

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 14:54:57 ; Search time 48.8955 Seconds  
(without alignments)  
69.343 Million cell updates/sec

Title: US-09-641-801-4  
Perfect score: 62  
Sequence: 1 LFFFLPVVNLP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	62	100.0	12	4	AAB72503	Aab72503 Colostrin
2	62	100.0	12	4	AAB59323	Aab59323 Ewe colos
3	62	100.0	12	4	AAB72249	Aab72249 Colostrin
4	62	100.0	12	4	AAB72535	Aab72535 Colostrin
5	62	100.0	12	5	AAO14580	Aao14580 Neural ce
6	62	100.0	12	5	AAM51039	Aam51039 Colostrin
7	62	100.0	12	5	AAE20231	Aae20231 Colostrin
8	62	100.0	14	4	AAB59353	Aab59353 Ewe colos
9	42	67.7	166	3	AAG33579	Aag33579 Arabidops

10	42	67.7	179	3	AAG33578	Aag33578 Arabidops
11	42	67.7	221	3	AAG33577	Aag33577 Arabidops
12	41	66.1	104	7	ADC96263	Adc96263 E. faeciu
13	41	66.1	126	6	ADA54377	Ada54377 Human pro
14	41	66.1	692	6	ABR53270	Abr53270 Protein s
15	40	64.5	733	6	ADA34955	Ada34955 Acinetoba
16	39	62.9	243	3	AAG54783	Aag54783 Arabidops
17	39	62.9	422	6	ABU21537	Abu21537 Protein e
18	39	62.9	710	5	ABB93859	Abb93859 Herbicida
19	39	62.9	750	3	AAG47557	Aag47557 Arabidops
20	38	61.3	113	6	ADB08914	Adb08914 Alloiococ
21	38	61.3	122	6	ADB08916	Adb08916 Alloiococ
22	38	61.3	338	6	ABM68047	Abm68047 Photorhab
23	38	61.3	678	5	ABB91909	Abb91909 Herbicida
24	38	61.3	708	3	AAG53600	Aag53600 Arabidops
25	38	61.3	709	3	AAG53599	Aag53599 Arabidops
26	38	61.3	709	5	ABB91016	Abb91016 Herbicida
27	38	61.3	728	5	ABB91071	Abb91071 Herbicida
28	38	61.3	738	3	AAG53598	Aag53598 Arabidops
29	37	59.7	42	2	AAW88741	Aaw88741 Secreted
30	37	59.7	42	4	ABB50508	Abb50508 Human sec
31	37	59.7	42	6	ABO44765	Abo44765 Novel hum
32	37	59.7	42	7	ABO26245	Abo26245 Human pro
33	37	59.7	67	4	ABG24626	Abg24626 Novel hum
34	37	59.7	72	7	ADC97216	Adc97216 E. faeciu
35	37	59.7	84	4	ABB03160	Abb03160 Human mus
36	37	59.7	84	6	ABU12454	Abu12454 Novel hum
37	37	59.7	103	2	AAW88536	Aaw88536 Secreted
38	37	59.7	103	4	ABB50303	Abb50303 Human sec
39	37	59.7	103	6	ABO44560	Abo44560 Novel hum
40	37	59.7	103	7	ABO26040	Abo26040 Human pro
41	37	59.7	141	4	ABG13456	Abg13456 Novel hum
42	37	59.7	159	4	ABG13434	Abg13434 Novel hum
43	37	59.7	163	4	ABG13496	Abg13496 Novel hum
44	37	59.7	288	4	ABG26638	Abg26638 Novel hum
45	37	59.7	296	4	ABG13294	Abg13294 Novel hum

# ALIGNMENTS

RESULT 1

AAB72503

ID AAB72503 standard; peptide; 12 AA.

XX

AC AAB72503;

XX

DT 09-MAY-2001 (first entry)

XX

DE Colostrinin peptide #4.

XX

KW Dermatological; oxidative stress regulator; colostrinin.

XX

OS Unidentified.

XX

PN WO200112650-A2.

XX

PD 22-FEB-2001.  
 XX  
 PF 17-AUG-2000; 2000WO-US022665.  
 XX  
 PR 17-AUG-1999; 99US-0149310P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Stanton GJ, Hughes TK, Boldogh I;  
 XX  
 DR WPI; 2001-218342/22.  
 XX  
 PT Modulating oxidative stress level in a cell, involves contacting the cell  
 PT with an oxidative stress regulator selected from colostrinin, its  
 PT constituent peptide, analog or their combinations.  
 XX  
 PS Claim 6; Page 25; 48pp; English.  
 XX  
 CC The present invention relates to a method for modulating the oxidative  
 CC stress level in a cell or a patient, comprising contacting the cell with,  
 CC or administering to the patient, an oxidative stress regulator selected  
 CC from colostrinin, or its constituent peptide (e.g. the present peptide),  
 CC to change the level of an oxidising species in the cell. The method can  
 CC be used to treat oxidative damage to skin, by decreasing or preventing an  
 CC increase in the level of damage to a biomolecule of the patient  
 XX  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 4; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00041;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
 |||||  
 Db 1 LFFFLPVVNVLP 12

# RESULT 2

AAB59323

ID AAB59323 standard; peptide; 12 AA.

XX

AC AAB59323;

XX

DT 21-MAR-2001 (first entry)

XX

DE Ewe colostrinin peptide fragment B-8.

XX

KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;  
 KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX

OS Ovis sp.

XX

PN WO200075173-A2.

XX

PD 14-DEC-2000.

XX

PF 02-JUN-2000; 2000WO-GB002128.

XX  
 PR 02-JUN-1999; 99GB-00012852.  
 XX  
 PA (REGE-) REGEN THERAPEUTICS PLC.  
 XX  
 PI Georgiades JA;  
 XX  
 DR WPI; 2001-071058/08.  
 XX  
 PT Peptides having an N-terminal amino acid sequence isolated from  
 PT colostrinin for treating e.g. disorders of the central nervous system and  
 PT immune system, viral and bacterial infections, and diseases characterized  
 PT by amyloid plaques.  
 XX  
 PS Claim 7; Page 27; 63pp; English.  
 XX  
 CC The present invention provides the sequences of a number of peptides  
 CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide  
 CC fragment of colostrum. These peptides can be used in the treatment of  
 CC central nervous system disorders such as senile dementia, Parkinson's  
 CC disease, Alzheimer's disease, psychosis and neurosis, immune system  
 CC disorders such as bacterial and viral infections, to improve the  
 CC development of a child's immune system, as a dietary supplement, and to  
 CC promote the dissolution of beta-amyloid plaques  
 XX  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 4; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00041;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
 |||||  
 Db 1 LFFFLPVVNVLP 12

# RESULT 3

AAB72249

ID AAB72249 standard; peptide; 12 AA.

XX

AC AAB72249;

XX

DT 14-MAY-2001 (first entry)

XX

DE Colostrinin derived cytokine inducing peptide SEQ ID 4.

XX

KW Colostrinin; immune response; cytokine; blood cell proliferation;  
 KW central nervous system disorder; neurological disorder; mental disorder;  
 KW dementia; neurodegenerative disease; Alzheimer's disease; psychosis;  
 KW neurosis; infection.

XX

OS Synthetic.

XX

PN WO200111937-A2.

XX

PD 22-FEB-2001.

XX

PF 17-AUG-2000; 2000WO-US022818.  
 XX  
 PR 17-AUG-1999; 99US-0149311P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (REGE-) REGEN THERAPEUTICS PLC.  
 XX  
 PI Stanton GJ, Hughes TK, Boldogh I, Georgiades J;  
 XX  
 DR WPI; 2001-202804/20.  
 XX  
 PT Inducing a cytokine and modulating an immune response, useful for  
 PT treating central nervous system diseases and bacterial and viral  
 PT infections, comprises administering colostrinin as an immunological  
 PT regulator.  
 XX  
 PS Claim 1; Page 34; 50pp; English.  
 XX  
 CC Sequences AAB72246 - AAB72275 represent peptides derived from clostrinin,  
 CC a proline rich polypeptide aggregate contained in colostrum. The peptides  
 CC have immune response modulatory activity, and are capable of inducing  
 CC cytokines. Colostrinin and its derived peptides are useful for inducing  
 CC cytokine production, for modulating an immunological response and for  
 CC inducing blood cell proliferation. The peptides are useful in the  
 CC treatment of disorders of the central nervous system, neurological  
 CC disorders, mental disorders, dementia, neurodegenerative diseases,  
 CC Alzheimer's disease, motor neurone disease, psychosis, neurosis, chronic  
 CC disorders of the immune system, bacterial and viral infections and  
 CC acquired immunological deficiencies  
 XX  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 4; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00041;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVL 12  
 |||||  
 Db 1 LFFFLPVVNVL 12

#### RESULT 4

AAB72535

ID AAB72535 standard; peptide; 12 AA.

XX

AC AAB72535;

XX

DT 09-MAY-2001 (first entry)

XX

DE Colostrinin peptide #4.

XX

KW Neuroprotective; neural cell differentiation regulator; colostrinin;  
 KW colostrum.

XX

OS Unidentified.

XX

PN WO200112651-A2.

XX  
 PD 22-FEB-2001.  
 XX  
 PF 17-AUG-2000; 2000WO-US022774.  
 XX  
 PR 17-AUG-1999; 99US-0149633P.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Boldogh I;  
 XX  
 DR WPI; 2001-226545/23.  
 XX  
 PT Use of colostrinin, its constituent peptide or analog as a neural cell  
 PT regulator, for promoting neural cell differentiation and treating damaged  
 PT neural cells in a patient.  
 XX  
 PS Claim 6; Page 21; 35pp; English.  
 XX  
 CC The present invention relates to a method for promoting neural cell  
 CC differentiation and treating damaged neural cells, using colostrinin and  
 CC colostrinin constituent peptides (e.g. the present peptide) as a neural  
 CC cell regulator. Colostrinin is a polypeptide complex found in colostrum  
 XX  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 4; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00041;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
 |||||  
 Db 1 LFFFLPVVNVLP 12

# RESULT 5

AAO14580

ID AAO14580 standard; peptide; 12 AA.

XX

AC AAO14580;

XX

DT 27-MAY-2002 (first entry)

XX

DE Neural cell regulatory colostrinin peptide 4.

XX

KW Neural cell differentiation; neural cell regulator; colostrinin peptide;  
 KW neural cell formation; proline-rich polypeptide aggregate; colostrum;  
 KW neural cell treatment.

XX

OS Unidentified.

XX

FH Key Location/Qualifiers

FT Modified-site 12

FT /note= "Optional C-terminal amide"

XX

PN WO200213851-A1.

XX

PD 21-FEB-2002.  
 XX  
 PF 17-AUG-2000; 2000WO-US022777.  
 XX  
 PR 17-AUG-2000; 2000WO-US022777.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Boldogh I, Stanton JG, Hughes TK;  
 XX  
 DR WPI; 2002-269152/31.  
 XX  
 PT Promoting cell differentiation in a patient involves use of blood cell  
 PT regulator selected from colostrinin, its constituent peptide and/or  
 PT analog.  
 XX  
 PS Claim 7; Page 21; 37pp; English.  
 XX  
 CC The invention comprises a method for promoting cell differentiation (e.g.  
 CC neural cell differentiation). The method involves contacting cells with a  
 CC neural cell regulator (i.e. a colostrinin peptide) in order to change the  
 CC cells in morphology to form neural cells. Colostrinin is a proline-rich  
 CC polypeptide aggregate that is present in colostrum. The method of the  
 CC invention is useful for promoting the differentiation of cells and for  
 CC treating damaged neural cells in a patient. The present amino acid  
 CC sequence represents a specifically claimed colostrinin peptide used in  
 CC the method of the invention  
 XX  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 5; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00041;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVL 12  
 |||||  
 Db 1 LFFFLPVVNVL 12

# RESULT 6

AAM51039

ID AAM51039 standard; peptide; 12 AA.

XX

AC AAM51039;

XX

DT 30-MAY-2002 (first entry)

XX

DE Colostrinin constituent peptide.

XX

KW Colostrinin; colostrum; immunomodulator; cardiovascular;  
 KW blood cell regulator; cytokine inducer; beta-casein; human.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Modified-site 12

FT /note= "optional C-terminal amidation"

XX  
 PN WO200213849-A1.  
 XX  
 PD 21-FEB-2002.  
 XX  
 PF 17-AUG-2000; 2000WO-US022775.  
 XX  
 PR 17-AUG-2000; 2000WO-US022775.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PA (REGE-) REGEN THERAPEUTICS PLC.  
 XX  
 PI Stanton GJ, Hughes TK, Boldogh I, Georgiades J;  
 XX  
 DR WPI; 2002-269150/31.  
 XX  
 PT Modulation of blood cell proliferation in a patient involves use of blood  
 PT cell regulator selected from colostrinin, its constituent peptide and/or  
 PT analog.  
 XX  
 PS Claim 1; Page 34; 54pp; English.  
 XX  
 CC The present sequence is that of a colostrinin constituent peptide that is  
 CC preferred for use as an immunological regulator and as a blood cell  
 CC regulator in claimed methods of the invention. It is classified as having  
 CC a beta-casein homologue precursor. Methods are claimed for: inducing a  
 CC cytokine in a cell by contact with an immunological regulator, where the  
 CC cell is present in a cell culture, a tissue, an organ or an organism, and  
 CC the cell is mammalian, including human; modulating an immune response in  
 CC a cell by contact with the immunological regulator under conditions  
 CC effective to induce a cytokine; modulating an immune response in a  
 CC patient by administering an immunological regulator under conditions  
 CC effective to induce a cytokine, where the immunological regulator is  
 CC administered topically or as part of a dietary supplement, and where the  
 CC immune response is specific or non specific, an interferon response or an  
 CC antibody response; modulating blood cell proliferation by contacting  
 CC blood cells with a blood cell regulator, where the blood cells are  
 CC present in a cell culture or an organism, are mammalian or human, and  
 CC where the blood cells are increased in number or differentiated; and a  
 CC method for modulating blood cell proliferation in a patent. A claimed  
 CC cytokine-inducing composition comprises a pharmaceutical carrier and an  
 CC active agent such as the present peptide. Cytokines induced by this  
 CC peptide in human leucocyte cultures include interferon-gamma, tumour  
 CC necrosis factor-alpha, interleukin-6 and interleukin-10  
 XX  
 SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 5; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.00041;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
 |||||  
 Db 1 LFFFLPVVNVLP 12

RESULT 7



AAE20231

ID AAE20231 standard; peptide; 12 AA.

XX

AC AAE20231;

XX

DT 18-JUN-2002 (first entry)

XX

DE Colostrinin constituent peptide #4.

XX

KW Blood cell regulator; colostrinin; constituent peptide; oxidative stress;

KW therapy; oxidative damage; skin; aging; wound healing; cell replacement;

KW tissue; organ; cosmetic procedure; repair; regeneration; preservation;

KW transplantation; implantation; dermatological; vulnerary.

XX

OS Unidentified.

XX

FH Key Location/Qualifiers

FT Modified-site 12

FT /note= "Optionally C-terminal amide"

XX

PN WO200213850-A1.

XX

PD 21-FEB-2002.

XX

PF 17-AUG-2000; 2000WO-US022776.

XX

PR 17-AUG-2000; 2000WO-US022776.

XX

PA (TEXA ) UNIV TEXAS SYSTEM.

XX

PI Stanton GJ, Hughes TK, Boldogh I;

XX

DR WPI; 2002-269151/31.

XX

PT Composition useful for the modulation of blood cell proliferation in a  
PT patient comprises a blood cell regulator selected from colostrinin, its  
PT constituent peptide and/or analog.

XX

PS Claim 6; Page 25; 51pp; English.

XX

CC The invention relates to a composition which comprises a blood cell  
CC regulator selected from colostrinin, its constituent peptide and/or  
CC analogue. The invention is used for modulating the oxidative stress level  
CC in a cell e.g. mammalian or human cell present in a cell culture, tissue,  
CC organ, or organism; or for treating oxidative damage to the skin of a  
CC patient e.g. animal or human; to modulate oxidative stress during/ after  
CC a premature birth or normal birth, preventing/delaying aging in a  
CC patient, enhancing wound healing, and the reduction of side effects of  
CC cosmetic procedures. The method changes the level of an oxidising species  
CC in the cell, such as decreases or prevents increase in the level of  
CC damage to a biomolecule of the patient selected from DNA, protein and/or  
CC lipid, compared to the same conditions when the oxidative stress  
CC regulator is not present. The modulation of oxidative stress results in  
CC enhanced repair, regeneration, and replacement of cells, tissues and  
CC organs (e.g. kidney, liver, pancreas, skin, and the other internal and  
CC external organs), as well as enhanced preservation of such organs for  
CC transplantation, implantation, or scientific research. The present

CC sequence is a colostrinin constituent peptide  
XX  
SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 5; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00041;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LFFFLPVVNVL 12  
| | | | | | | | | |  
Db 1 LFFFLPVVNVL 12

RESULT 8

AAB59353

ID AAB59353 standard; peptide; 14 AA.

XX

AC AAB59353;

XX

DT 21-MAR-2001 (first entry)

XX

DE Ewe colostrinin peptide fragment derived sequence #13.

XX

KW Sheep; colostrinin; proline rich polypeptide; colostrum; immune disorder;  
KW central nervous system disorder; dietary supplement; beta-amyloid plaque.

XX

OS Ovis sp.

XX

PN WO200075173-A2.

XX

PD 14-DEC-2000.

XX

PF 02-JUN-2000; 2000WO-GB002128.

XX

PR 02-JUN-1999; 99GB-00012852.

XX

PA (REGE-) REGEN THERAPEUTICS PLC.

XX

PI Georgiades JA;

XX

DR WPI; 2001-071058/08.

XX

PT Peptides having an N-terminal amino acid sequence isolated from  
PT colostrinin for treating e.g. disorders of the central nervous system and  
PT immune system, viral and bacterial infections, and diseases characterized  
PT by amyloid plaques.

XX

PS Claim 8; Page 27; 63pp; English.

XX

CC The present invention provides the sequences of a number of peptides  
CC found in ewe's colostrinin. Colostrinin is the proline-rich polypeptide  
CC fragment of colostrum. These peptides can be used in the treatment of  
CC central nervous system disorders such as senile dementia, Parkinson's  
CC disease, Alzheimer's disease, psychosis and neurosis, immune system  
CC disorders such as bacterial and viral infections, to improve the  
CC development of a child's immune system, as a dietary supplement, and to  
CC promote the dissolution of beta-amyloid plaques

XX

SQ Sequence 14 AA;

Query Match 100.0%; Score 62; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00049;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
| | | | | | | | | |  
Db 2 LFFFLPVVNVLP 13

RESULT 9

AAG33579

ID AAG33579 standard; protein; 166 AA.

XX

AC AAG33579;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 40711.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.

PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.

PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 67.7%; Score 42; DB 3; Length 166;  
 Best Local Similarity 70.0%; Pred. No. 19;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVL 11  
 :|||||:| |  
 Db 135 YFFLPVINXL 144

RESULT 10  
 AAG33578  
 ID AAG33578 standard; protein; 179 AA.  
 XX

AC AAG33578;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 40710.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.

PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.



PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.

PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 67.7%; Score 42; DB 3; Length 179;  
 Best Local Similarity 70.0%; Pred. No. 21;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVL 11  
 :|||||:| |  
 Db 148 YFFLPVINXL 157

# RESULT 11

AAG33577

ID AAG33577 standard; protein; 221 AA.

XX

AC AAG33577;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 40709.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.

PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.

PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.

PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 67.7%; Score 42; DB 3; Length 221;  
 Best Local Similarity 70.0%; Pred. No. 26;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFFLPVVNVL 11  
 :||||:| |  
 Db 190 YFFLPVINXL 199

# RESULT 12

ADC96263

ID ADC96263 standard; protein; 104 AA.

XX

AC ADC96263;

XX

DT 01-JAN-2004 (first entry)

XX

DE E. faecium protein sequence SEQ ID 5890.

XX

KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;  
 KW abdominal-pelvic infection.

XX

OS Enterococcus faecium.

XX

PN US6583275-B1.

XX

PD 24-JUN-2003.

XX

PF 30-JUN-1998; 98US-00107532.

XX

PR 02-JUL-1997; 97US-0051571P.

PR 14-MAY-1998; 98US-0085598P.

XX

PA (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Doucette-Stamm LA, Bush D;

XX

DR WPI; 2003-799836/75.

DR N-PSDB; ADC92609.

XX

PT New isolated nucleic acid derived from Enterococcus faecium encoding an  
 PT Enterococcus faecium polypeptide useful for detection, prevention and  
 PT treatment of a pathological condition resulting from a bacterial  
 PT infection.

XX

PS Example 1; SEQ ID NO 5890; 243pp; English.

XX

CC The invention relates to an isolated nucleic acid derived from  
CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having  
CC one of 10 fully defined sequences given in the (or comprising 40  
CC sequential nucleotides chosen from any of the nucleic acids, its  
CC complement or sequences hybridising to it). Also included are a  
CC recombinant vector comprising the nucleic acid operably linked to  
CC transcription regulatory element, a cell comprising the vector and a  
CC single-stranded probe comprising the nucleic acid. The nucleic acids are  
CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.  
CC The nucleic acids is useful for diagnosing pathological conditions  
CC resulting from E. faecium bacterial infection (e.g. urinary tract  
CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic  
CC infection) and for screening drugs such as agonists and antagonists. The  
CC nucleic acid is useful for recombinant production of Candida albicans -  
CC derived peptides or antisense polypeptides. Pharmaceutical compositions  
CC and vaccines containing the nucleic acid are useful for preventing or  
CC treating Enterococcus faecium infections. The present sequence represents  
CC one if the disclosed E. faecium proteins.

XX

SQ Sequence 104 AA;

Query Match 66.1%; Score 41; DB 7; Length 104;  
Best Local Similarity 66.7%; Pred. No. 17;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFFLPVVNV 10

|||:|::||

Db 91 FFFIPLINV 99

#### RESULT 13

ADA54377

ID ADA54377 standard; protein; 126 AA.

XX

AC ADA54377;

XX

DT 20-NOV-2003 (first entry)

XX

DE Human protein, SEQ ID 1945.

XX

KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;

KW Gene Therapy; human; secretory protein; membrane proteins; cancer;

KW inflammatory disease; osteoporosis; neurological disease.

XX

OS Homo sapiens.

XX

PN EP1293569-A2.

XX

PD 19-MAR-2003.

XX

PF 21-MAR-2002; 2002EP-00006586.

XX

PR 14-SEP-2001; 2001JP-00328381.

PR 24-JAN-2002; 2002US-0350435P.

XX

PA (HELI-) HELIX RES INST.  
PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX  
DR WPI; 2003-395539/38.  
DR N-PSDB; ADA52738.  
XX  
PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
PT and/or membrane proteins, useful for developing medicines for diseases in  
PT which the gene is involved, or as target molecules for gene therapy.  
XX  
PS Claim 14; SEQ ID NO 1945; 205pp; English.  
XX  
CC The present invention relates to novel human secretory or membrane  
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
CC ADA54071). The coding sequences are useful in the gene therapy of  
CC diseases caused by abnormalities of the proteins, e.g. cancer,  
CC inflammatory diseases, osteoporosis or neurological disease.  
XX  
SQ Sequence 126 AA;

Query Match 66.1%; Score 41; DB 6; Length 126;  
Best Local Similarity 72.7%; Pred. No. 21;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12  
| | | | | | : | |  
Db 35 FFFLPPVSSLP 45

RESULT 14  
ABR53270  
ID ABR53270 standard; protein; 692 AA.  
XX  
AC ABR53270;  
XX  
DT 20-JUN-2003 (first entry)  
XX  
DE Protein sequence #SEQ ID 1405.  
XX  
KW Multiprotein complex; eukaryote; drug target; diagnosis.  
XX  
OS Saccharomyces cerevisiae.  
XX  
PN EP1258494-A1.  
XX  
PD 20-NOV-2002.  
XX  
PF 20-DEC-2001; 2001EP-00130253.  
XX  
PR 15-MAY-2001; 2001EP-00111774.  
XX  
PA (CELL-) CELLZOME AG.  
XX



PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;  
 PI Marzioch M, Schultz JD, Superti-Furga GD;  
 XX  
 DR WPI; 2003-250078/25.  
 DR N-PSDB; ACC61312.  
 XX  
 PT New isolated protein complexes useful for diagnosing a disease or  
 PT disorder, or as a target for an active agent of a pharmaceutical,  
 PT preferably a drug target in the treatment or prevention of disease or  
 PT disorder.  
 XX  
 PS Disclosure; SEQ ID NO 1405; 17pp + Sequence Listing; English.  
 XX  
 CC The invention relates to multiprotein complexes from eukaryotes. Proteins  
 CC of the invention and DNA sequences encoding them are given in records  
 CC ABR52568-ABR53903 and ACC60610-ACC61944 respectively. The complexes are  
 CC obtainable by using a protein as a bait and isolating the set of proteins  
 CC which is attached thereto from cells. Such protein complexes may comprise  
 CC up to 30 distinct proteins. Protein complexes of the invention are useful  
 CC for diagnosing a disease or disorder, or as a target for an active agent  
 CC of a pharmaceutical, preferably a drug target in the treatment or  
 CC prevention of a disease or disorder. Note: The sequence data for this  
 CC patent is not represented in the printed specification, but is based on  
 CC sequence information supplied by the European Patent Office. The complete  
 CC document is available on CD-ROM  
 XX  
 SQ Sequence 692 AA;

Query Match 66.1%; Score 41; DB 6; Length 692;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 LFFFLPVVNVLP 12  
 ||||: : ||||  
 Db 242 LFFFIILENVLP 253

# RESULT 15

ADA34955

ID ADA34955 standard; protein; 733 AA.

XX

AC ADA34955;

XX

DT 20-NOV-2003 (first entry)

XX

DE Acinetobacter baumannii protein #2116.

XX

KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;  
 KW plant biocontrol agent.

XX

OS Acinetobacter baumannii.

XX

PN US6562958-B1.

XX

PD 13-MAY-2003.

XX

PF 04-JUN-1999; 99US-00328352.

XX  
 PR 09-JUN-1998; 98US-0088701P.  
 XX  
 PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX  
 PI Breton G, Bush D;  
 XX  
 DR WPI; 2003-576092/54.  
 DR N-PSDB; ADA30829.  
 XX  
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents  
 PT for diagnosing a bacterial disease, as components of antibacterial  
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for  
 PT plants.  
 XX  
 PS Example; SEQ ID NO 6242; 328pp; English.  
 XX  
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.  
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents  
 CC for diagnosing a bacterial disease, as components of antibacterial  
 CC vaccines, as targets for antibacterial drugs, to detect the presence of  
 CC A. baumannii and other Acinetobacter species in a sample, in screening  
 CC compounds for the ability to interfere with the A. baumannii life cycle  
 CC or to inhibit A. baumannii infection, and as biocontrol agents for  
 CC plants. The present sequence represents the amino acid sequence of an A.  
 CC baumannii protein.  
 XX  
 SQ Sequence 733 AA;

Query Match 64.5%; Score 40; DB 6; Length 733;  
 Best Local Similarity 41.7%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
 : ||:|::|::|  
 Db 340 ILFFVPLNMNIP 351

Search completed: August 24, 2004, 15:42:24  
 Job time : 52.8955 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:33:13 ; Search time 13.1642 Seconds  
(without alignments)  
47.060 Million cell updates/sec

Title: US-09-641-801-4  
Perfect score: 62  
Sequence: 1 LFFFLPVVNVLP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	62	100.0	12	4	US-09-641-803-4	Sequence 4, Appli
2	41	66.1	104	4	US-09-107-532A-5890	Sequence 5890, Ap
3	40	64.5	733	4	US-09-328-352-6242	Sequence 6242, Ap
4	39	62.9	528	4	US-09-252-991A-23368	Sequence 23368, A
5	38	61.3	483	4	US-09-489-039A-7429	Sequence 7429, Ap
6	37	59.7	42	4	US-09-205-258-456	Sequence 456, App
7	37	59.7	72	4	US-09-107-532A-6843	Sequence 6843, Ap
8	37	59.7	103	4	US-09-205-258-251	Sequence 251, App
9	37	59.7	405	4	US-09-489-039A-12853	Sequence 12853, A
10	36	58.1	48	4	US-09-071-035-196	Sequence 196, App
11	36	58.1	109	4	US-09-071-035-194	Sequence 194, App

12	36	58.1	123	4	US-09-134-000C-4445	Sequence 4445, Ap
13	36	58.1	137	4	US-09-489-039A-11180	Sequence 11180, A
14	36	58.1	307	2	US-08-782-760-6	Sequence 6, Appli
15	36	58.1	307	5	PCT-US96-00995-6	Sequence 6, Appli
16	36	58.1	396	1	US-07-649-591B-4	Sequence 4, Appli
17	36	58.1	396	1	US-08-277-540-4	Sequence 4, Appli
18	36	58.1	396	1	US-08-430-787A-4	Sequence 4, Appli
19	36	58.1	424	4	US-09-134-000C-5836	Sequence 5836, Ap
20	35	56.5	204	4	US-09-134-000C-3659	Sequence 3659, Ap
21	35	56.5	227	4	US-09-904-615-126	Sequence 126, App
22	35	56.5	271	3	US-09-077-675A-12	Sequence 12, Appl
23	35	56.5	271	4	US-09-077-674-12	Sequence 12, Appl
24	35	56.5	289	3	US-09-077-675A-10	Sequence 10, Appl
25	35	56.5	289	4	US-09-077-674-10	Sequence 10, Appl
26	35	56.5	302	3	US-09-077-675A-7	Sequence 7, Appli
27	35	56.5	302	4	US-09-077-674-7	Sequence 7, Appli
28	35	56.5	346	4	US-09-543-681A-4493	Sequence 4493, Ap
29	35	56.5	361	3	US-09-077-675A-8	Sequence 8, Appli
30	35	56.5	361	4	US-09-077-674-8	Sequence 8, Appli
31	35	56.5	366	3	US-09-077-675A-13	Sequence 13, Appl
32	35	56.5	366	4	US-09-077-674-13	Sequence 13, Appl
33	35	56.5	366	4	US-09-170-496D-88	Sequence 88, Appl
34	35	56.5	366	4	US-09-170-496D-210	Sequence 210, App
35	35	56.5	366	4	US-09-743-742B-7	Sequence 7, Appli
36	35	56.5	366	4	US-09-762-661A-5	Sequence 5, Appli
37	35	56.5	366	4	US-09-364-425B-45	Sequence 45, Appl
38	35	56.5	366	4	US-09-743-475-4	Sequence 4, Appli
39	35	56.5	398	4	US-09-107-532A-4954	Sequence 4954, Ap
40	35	56.5	423	1	US-07-649-591B-3	Sequence 3, Appli
41	35	56.5	423	1	US-08-277-540-3	Sequence 3, Appli
42	35	56.5	423	1	US-08-430-787A-3	Sequence 3, Appli
43	35	56.5	423	2	US-08-869-057-2	Sequence 2, Appli
44	35	56.5	423	4	US-09-813-133A-4	Sequence 4, Appli
45	35	56.5	473	4	US-09-543-681A-7980	Sequence 7980, Ap

# ALIGNMENTS

## RESULT 1

US-09-641-803-4

; Sequence 4, Application US/09641803

; Patent No. 6500798

; GENERAL INFORMATION:

; APPLICANT: STANTON, G. John

; APPLICANT: HUGHES, Thomas K.

; APPLICANT: BOLDOGH, Istvan

; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND

; TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS

; FILE REFERENCE: 265.00220101

; CURRENT APPLICATION NUMBER: US/09/641,803

; CURRENT FILING DATE: 2000-08-17

; PRIOR APPLICATION NUMBER: 60/149,310

; PRIOR FILING DATE: 1999-08-17

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: peptide  
US-09-641-803-4

Query Match 100.0%; Score 62; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
| | | | | | | | | |  
Db 1 LFFFLPVVNVLP 12

RESULT 2

US-09-107-532A-5890

; Sequence 5890, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND

THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 5890:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 104 amino acids

;           TYPE: amino acid  
;           TOPOLOGY: linear  
;       MOLECULE TYPE: protein  
;       HYPOTHETICAL: YES  
;       ORIGINAL SOURCE:  
;           ORGANISM: Enterococcus faecium  
;       FEATURE:  
;           NAME/KEY: misc\_feature  
;           LOCATION: (B) LOCATION 1...104  
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5890:  
US-09-107-532A-5890

Query Match                   66.1%;   Score 41;   DB 4;   Length 104;  
Best Local Similarity   66.7%;   Pred. No. 7.9;  
Matches       6;   Conservative       3;   Mismatches       0;   Indels       0;   Gaps       0;

Qy           2 FFFLPVVNV 10  
             |||:|::||  
Db           91 FFFIPLINV 99

RESULT 3

US-09-328-352-6242  
; Sequence 6242, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
;   APPLICANT: Gary L. Breton et al.  
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ACINETOBACTER  
;   TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
;   FILE REFERENCE: GTC99-03PA  
;   CURRENT APPLICATION NUMBER: US/09/328,352  
;   CURRENT FILING DATE: 1999-06-04  
;   NUMBER OF SEQ ID NOS: 8252  
;   SEQ ID NO 6242  
;   LENGTH: 733  
;   TYPE: PRT  
;   ORGANISM: Acinetobacter baumannii  
US-09-328-352-6242

Query Match                   64.5%;   Score 40;   DB 4;   Length 733;  
Best Local Similarity   41.7%;   Pred. No. 88;  
Matches       5;   Conservative       6;   Mismatches       1;   Indels       0;   Gaps       0;

Qy           1 LFFFLPVVNVLP 12  
             : ||:|::|::|  
Db           340 ILFFVPLNMNIP 351

RESULT 4

US-09-252-991A-23368  
; Sequence 23368, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
;   APPLICANT: Marc J. Rubenfield et al.  
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23368  
; LENGTH: 528  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23368

Query Match 62.9%; Score 39; DB 4; Length 528;  
Best Local Similarity 50.0%; Pred. No. 91;  
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVL 11  
| | | : | : : : |  
Db 395 FFFMPILSIL 404

RESULT 5

US-09-489-039A-7429  
; Sequence 7429, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 7429  
; LENGTH: 483  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-7429

Query Match 61.3%; Score 38; DB 4; Length 483;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVL 11  
: | | | : | | |  
Db 130 YLFLPMINVL 139

RESULT 6

US-09-205-258-456  
; Sequence 456, Application US/09205258

; Patent No. 6525174  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: PZ007P1  
; CURRENT APPLICATION NUMBER: US/09/205,258  
; CURRENT FILING DATE: 1998-12-04  
; EARLIER APPLICATION NUMBER: PCT/US98/11422  
; EARLIER FILING DATE: 1998-06-04  
; EARLIER APPLICATION NUMBER: 60/048,885  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,375  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,881  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,880  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,020  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,876  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,895  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,884  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,894  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,971  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,882  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,899  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,893  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,900  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,901  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,892  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,915  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,019  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,970  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,972  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,916  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,373  
; EARLIER FILING DATE: 1997-06-06



```

; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 456
;   LENGTH: 42
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (42)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-205-258-456

```

```

Query Match          59.7%;   Score 37;   DB 4;   Length 42;
Best Local Similarity 75.0%;   Pred. No. 14;
Matches      6;   Conservative    2;   Mismatches    0;   Indels      0;   Gaps      0;

```

```

Qy      1 LFFFLPVV 8
        |||||::
Db      21 LFFFLPLI 28

```

```

RESULT 7
US-09-107-532A-6843
; Sequence 6843, Application US/09107532A
; Patent No. 6583275
;   GENERAL INFORMATION:
;       APPLICANT: Lynn A Doucette-Stamm   and David Bush

```

```

;      TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                          ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND
THERAPEUTICS
;      NUMBER OF SEQUENCES: 7310
;      CORRESPONDENCE ADDRESS:
;          ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;          STREET: 100 Beaver Street
;          CITY: Waltham
;          STATE: Massachusetts
;          COUNTRY: USA
;          ZIP: 02354
;      COMPUTER READABLE FORM:
;          MEDIUM TYPE: CD-ROM ISO9660
;          COMPUTER: PC
;          OPERATING SYSTEM: <Unknown>
;          SOFTWARE: ASCII
;      CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: US/09/107,532A
;          FILING DATE: 30-Jun-1998
;      PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: 60/085,598
;          FILING DATE: 14 May 1998
;          APPLICATION NUMBER: 60/051571
;          FILING DATE: July 2, 1997
;      ATTORNEY/AGENT INFORMATION:
;          NAME: Ariniello, Pamela Deneke
;          REGISTRATION NUMBER: 40,489
;          REFERENCE/DOCKET NUMBER: GTC-012
;      TELECOMMUNICATION INFORMATION:
;          TELEPHONE: (781)893-5007
;          TELEFAX: (781)893-8277
;      INFORMATION FOR SEQ ID NO: 6843:
;          SEQUENCE CHARACTERISTICS:
;              LENGTH: 72 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;      MOLECULE TYPE: protein
;      HYPOTHETICAL: YES
;      ORIGINAL SOURCE:
;          ORGANISM: Enterococcus faecium
;      FEATURE:
;          NAME/KEY: misc_feature
;          LOCATION: (B) LOCATION 1...72
;      SEQUENCE DESCRIPTION: SEQ ID NO: 6843:
US-09-107-532A-6843

```

```

Query Match          59.7%;  Score 37;  DB 4;  Length 72;
Best Local Similarity 54.5%;  Pred. No. 24;
Matches      6;  Conservative      2;  Mismatches      3;  Indels      0;  Gaps      0;

```

```

Qy      2 FFFLPVVNVLP 12
        || ||:| : |
Db      45 FFTLPIVKIYP 55

```

```

RESULT 8
US-09-205-258-251

```

; Sequence 251, Application US/09205258  
; Patent No. 6525174  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: PZ007P1  
; CURRENT APPLICATION NUMBER: US/09/205,258  
; CURRENT FILING DATE: 1998-12-04  
; EARLIER APPLICATION NUMBER: PCT/US98/11422  
; EARLIER FILING DATE: 1998-06-04  
; EARLIER APPLICATION NUMBER: 60/048,885  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,375  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,881  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,880  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,020  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,876  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,895  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,884  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,894  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,971  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,882  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,899  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,893  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,900  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,901  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,892  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,915  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,019  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,970  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,972  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,916  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,373

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 251
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (103)
; OTHER INFORMATION: Xaa equals stop translation
US-09-205-258-251

```

```

Query Match          59.7%;  Score 37;  DB 4;  Length 103;
Best Local Similarity 75.0%;  Pred. No. 35;
Matches      6;  Conservative      2;  Mismatches      0;  Indels      0;  Gaps      0;

```

```

Qy      1 LFFFLPVV 8
        |||||::
Db      82 LFFFLPLI 89

```

```

RESULT 9
US-09-489-039A-12853
; Sequence 12853, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

```

; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 12853  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12853

Query Match 59.7%; Score 37; DB 4; Length 405;  
Best Local Similarity 70.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FFLPVVNVLP 12  
|| | || ||  
Db 303 FFRPAVNFLP 312

RESULT 10

US-09-071-035-196

; Sequence 196, Application US/09071035  
; Patent No. 6448043  
; GENERAL INFORMATION:  
; APPLICANT: Gil H. Choi  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 496  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071,035  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: A. Anders Brookes  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB369P2  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 196:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 48 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-071-035-196

Query Match 58.1%; Score 36; DB 4; Length 48;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 PVNVLP 12  
| | | | |  
Db 24 PVNVLP 30

RESULT 11

US-09-071-035-194  
; Sequence 194, Application US/09071035  
; Patent No. 6448043  
; GENERAL INFORMATION:  
; APPLICANT: Gil H. Choi  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 496  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071,035  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: A. Anders Brookes  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB369P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 194:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 109 amino acids

; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-071-035-194

Query Match 58.1%; Score 36; DB 4; Length 109;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 PVVNVLP 12  
| | | | | | |  
Db 48 PVVNVLP 54

RESULT 12

US-09-134-000C-4445

; Sequence 4445, Application US/09134000C

; Patent No. 6617156

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 032796-032

; CURRENT APPLICATION NUMBER: US/09/134,000C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/055,778

; PRIOR FILING DATE: 1997-08-15

; NUMBER OF SEQ ID NOS: 6812

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4445

; LENGTH: 123

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-09-134-000C-4445

Query Match 58.1%; Score 36; DB 4; Length 123;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 PVVNVLP 12  
| | | | | | |  
Db 62 PVVNVLP 68

RESULT 13

US-09-489-039A-11180

; Sequence 11180, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
KLEBSIELLA

; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 11180  
; LENGTH: 137  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-11180

Query Match 58.1%; Score 36; DB 4; Length 137;  
Best Local Similarity 45.5%; Pred. No. 69;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVL 11  
| ||:||:::  
Db 123 LLFFIPVLSII 133

RESULT 14

US-08-782-760-6

; Sequence 6, Application US/08782760  
; Patent No. 5948668  
; GENERAL INFORMATION:  
; APPLICANT: Hartman, Jacob  
; APPLICANT: Fulga, Netta  
; APPLICANT: Mendelovitch, Simona  
; APPLICANT: Gorecki, Marian  
; TITLE OF INVENTION: PRODUCTION OF ENZYMATICALLY ACTIVE  
; TITLE OF INVENTION: CARBOXYPEPTIDASE B  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/782,760  
; FILING DATE: 13-JAN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/378,233  
; FILING DATE: 25-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0336/43847  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 6:



; SEQUENCE CHARACTERISTICS:  
; LENGTH: 307 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-782-760-6

Query Match 58.1%; Score 36; DB 2; Length 307;  
Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVVNV 10  
|: |||||:  
Db 102 FYVLPVVNI 110

RESULT 15

PCT-US96-00995-6

; Sequence 6, Application PC/TUS9600995  
; GENERAL INFORMATION:  
; APPLICANT: Bio-Technology General Corp.  
; TITLE OF INVENTION: PRODUCTION OF ENZYMATICALLY ACTIVE  
; TITLE OF INVENTION: CARBOXYPEPTIDASE B  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/00995  
; FILING DATE: 25-JAN-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0336/43847-A-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 308 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US96-00995-6

Query Match 58.1%; Score 36; DB 5; Length 307;  
Best Local Similarity 66.7%; Pred. No. 1.6e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 FFFLPVVNV 10  
|: ||||:  
Db 102 FYVLPVUNI 110

Search completed: August 24, 2004, 15:55:14  
Job time : 14.1642 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:26:28 ; Search time 11.6418 Seconds  
 (without alignments)  
 99.151 Million cell updates/sec

Title: US-09-641-801-4  
 Perfect score: 62  
 Sequence: 1 LFFFLPVVNLP 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : PIR\_78:\*  
 1: pirl:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result	%					Description
	No.	Score	Query Match	Length	ID	
1	48	77.4	469	2	AB3519	enterobactin synth
2	45	72.6	144	2	D90125	hypothetical prote
3	44	71.0	62	2	AI2652	hypothetical prote
4	41	66.1	66	2	S49114	hypothetical prote
5	41	66.1	488	2	D70420	NADH2 dehydrogenas
6	41	66.1	692	2	S61200	probable membrane
7	39	62.9	256	2	I49846	spa29 protein - Sh
8	39	62.9	359	2	T21247	hypothetical prote
9	39	62.9	391	2	S74688	hypothetical prote
10	39	62.9	471	2	F72257	lipopolysaccharide
11	39	62.9	509	2	E83002	drug efflux transp
12	39	62.9	537	2	G31277	quinat transport
13	39	62.9	710	2	T52573	cyclic nucleotide

14	38.5	62.1	364	2	AD2302	hypothetical prote
15	38	61.3	221	2	AF3483	heme exporter prot
16	38	61.3	296	2	AC2937	hypothetical prote
17	38	61.3	296	2	D98345	SN-glycerol-3-phos
18	38	61.3	461	2	H85755	probable amino aci
19	38	61.3	479	2	A90863	probable amino aci
20	38	61.3	479	2	C64878	probable amino aci
21	38	61.3	532	2	G82872	cardiolipin syntha
22	38	61.3	678	2	G84682	hypothetical prote
23	38	61.3	738	2	E86294	hypothetical prote
24	38	61.3	746	2	H86330	probable cyclic nu
25	37	59.7	218	2	T15420	hypothetical prote
26	37	59.7	263	2	S37309	spaR protein - Sal
27	37	59.7	263	2	AD0851	secretory protein
28	37	59.7	546	2	G71348	probable apolipoppr
29	37	59.7	851	2	T31520	hypothetical prote
30	36	58.1	118	2	S76608	hypothetical prote
31	36	58.1	201	2	T33593	hypothetical prote
32	36	58.1	249	2	G64470	sulfate/thiosulfat
33	36	58.1	252	2	A69286	polysaccharide ABC
34	36	58.1	302	2	E72402	phospho-N-acetylm
35	36	58.1	306	1	CPBOB	carboxypeptidase B
36	36	58.1	310	2	AB0979	hypothetical prote
37	36	58.1	310	2	F90011	conserved hypothet
38	36	58.1	327	1	HLHUCD	T-cell surface gly
39	36	58.1	415	2	A32129	carboxypeptidase B
40	36	58.1	438	2	H82280	probable vibriobac
41	36	58.1	488	2	T30602	hypothetical prote
42	36	58.1	501	2	T52170	cytochrome P450 mo
43	36	58.1	505	2	T02898	hypothetical prote
44	36	58.1	555	2	B72341	uridine kinase-rel
45	36	58.1	689	2	D85013	hypothetical prote

## ALIGNMENTS

### RESULT 1

AB3519

enterobactin synthetase component F [imported] - *Brucella melitensis* (strain 16M)

C;Species: *Brucella melitensis*

C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002

C;Accession: AB3519

R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N.; Anderson, I.; Bhattacharyya, A.; Lykidis, A.; Reznik, G.; Jablonski, L.; Larsen, N.; D'Souza, M.; Bernal, A.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesson, J.J.; Haselkorn, R.; Kyrpides, N.; Overbeek, R.

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*.

A;Reference number: AD3252; PMID:11756688

A;Accession: AB3519

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-469 <KUR>

A;Cross-references: GB:AE008918; PIDN:AAL53317.1; PID:g17984203; GSPDB:GN00191  
A;Experimental source: strain 16M  
C;Genetics:  
A;Gene: BMEII0076  
A;Map position: II

Query Match 77.4%; Score 48; DB 2; Length 469;  
Best Local Similarity 72.7%; Pred. No. 0.94;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12  
||| |::|||  
Db 374 FFFSPLINVLP 384

RESULT 2

D90125

hypothetical protein orf144 [imported] - Guillardia theta nucleomorph

C;Species: nucleomorph Guillardia theta

A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001

C;Accession: D90125

R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reith, M.; Cavalier-Smith, T.; Maier, U.G.

Nature 410, 1091-1096, 2001

A;Title: The highly reduced genome of an enslaved algal nucleus.

A;Reference number: A99082; MUID:11323671; PMID:11323671

A;Accession: D90125

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-144 <DOU>

A;Cross-references: GB:AF083031; NID:g13794319; PIDN:AAK39696.1; GSPDB:GN00152

C;Genetics:

A;Gene: orf144

A;Map position: 3

A;Genome: nucleomorph

C;Keywords: nucleomorph

Query Match 72.6%; Score 45; DB 2; Length 144;  
Best Local Similarity 66.7%; Pred. No. 0.97;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
:|||| ||:|  
Db 24 IFFFLKKNILP 35

RESULT 3

AI2652

hypothetical protein Atu0623 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002

C;Accession: AI2652

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Kitajima, J.P.; Okura, V.K.; Almeida Jr., N.F.; Zhou, Y.; Bovee Sr., D.; Chapman, P.; Clendenning, J.; Deatherage, G.; Gillet, W.; Grant, C.;

Guenthner, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClelland, E.; Palmieri, A.; Raymond, C.; Rouse, G.; Saenphimmachak, C.; Wu, Z.; Gordon, D.; Eisen, J.A.; Paulsen, I.; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, B.; Liao, L.; Kim, S.; Hendrick, C.; Zhao, Z.; Dolan, M.; Tingey, S.V.; Tomb, J.; Gordon, M.P.; Olson, M.V.; Nester, E.W.  
 A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
 A;Reference number: AB2577; MUID:21608550; PMID:11743193  
 A;Accession: AI2652  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-62 <KUR>  
 A;Cross-references: GB:AE008688; PIDN:AAL41639.1; PID:g17738979; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: Atu0623  
 A;Map position: circular chromosome

Query Match 71.0%; Score 44; DB 2; Length 62;  
 Best Local Similarity 72.7%; Pred. No. 0.62;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12  
 |:|:|:|:|  
 Db 5 FYSLPVMNVLP 15

#### RESULT 4

S49114

hypothetical protein - yeast (*Williopsis suaveolens*) mitochondrion (fragment)

C;Species: mitochondrion *Williopsis suaveolens*

C;Date: 16-Feb-1995 #sequence\_revision 26-May-1995 #text\_change 07-Dec-1999

C;Accession: S49114

R;Nosek, J.

submitted to the EMBL Data Library, January 1994

A;Reference number: S49114

A;Accession: S49114

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-66 <NOS>

A;Cross-references: EMBL:X77238; NID:g509359; PID:g509360

C;Genetics:

A;Genome: mitochondrion

A;Genetic code: SGC2

C;Keywords: mitochondrion

Query Match 66.1%; Score 41; DB 2; Length 66;  
 Best Local Similarity 50.0%; Pred. No. 2.2;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
 |:|:|:|:|:|  
 Db 37 LFFFIMIIGVMP 48

# RESULT 5

D70420

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain nuoN2 - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 03-Jun-2002

C;Accession: D70420

R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Overbeek, R.; Snead, M.A.; Keller, M.; Aujay, M.; Huber, R.; Feldman, R.A.; Short, J.M.; Olson, G.J.; Swanson, R.V.

Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: D70420

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-488 <AQF>

A;Cross-references: GB:AE000737; NID:g2983782; PIDN:AAC07354.1; PID:g2983796; GB:AE000657

A;Experimental source: strain VF5

C;Genetics:

A;Gene: nuoN2

C;Superfamily: NADH dehydrogenase (ubiquinone) chain 2

C;Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 66.1%; Score 41; DB 2; Length 488;

Best Local Similarity 58.3%; Pred. No. 17;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVL 12

| ||:|:| |:|

Db 255 LAFFIPLVRVMP 266

# RESULT 6

S61200

probable membrane protein YDR314c - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein D9740.21

C;Species: Saccharomyces cerevisiae

C;Date: 23-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 19-Apr-2002

C;Accession: S61200

R;Ding, H.

submitted to the EMBL Data Library, June 1995

A;Description: The sequence of S. cerevisiae cosmid 9740.

A;Reference number: S61160

A;Accession: S61200

A;Molecule type: DNA

A;Residues: 1-692 <DIN>

A;Cross-references: EMBL:U28374; NID:g849207; PIDN:AAB64750.1; PID:g849228; GSPDB:GN00004; MIPS:YDR314c

C;Genetics:

A;Gene: MIPS:YDR314c

A;Cross-references: SGD:S0002722

A;Map position: 4R

C;Superfamily: yeast probable membrane protein YDR314c

C;Keywords: transmembrane protein

F;94-110/Domain: transmembrane #status predicted <TM1>

F;239-255/Domain: transmembrane #status predicted <TM2>

Query Match 66.1%; Score 41; DB 2; Length 692;  
Best Local Similarity 66.7%; Pred. No. 24;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNLP 12  
||||: : ||||  
Db 242 LFFFIILENVLP 253

RESULT 7

I49846

spa29 protein - Shigella flexneri plasmid pMYSH6000

C;Species: Shigella flexneri

C;Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 28-Jul-2000

C;Accession: I49846

R;Sasakawa, C.; Komatsu, K.; Tobe, T.; Suzuki, T.; Yoshikawa, M.

J. Bacteriol. 175, 2334-2346, 1993

A;Title: Eight genes in region 5 that form an operon are essential for invasion of epithelial cells by Shigella flexneri 2a.

A;Reference number: A49846; MUID:93224456; PMID:8385666

A;Accession: I49846

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-256 <SAS>

A;Cross-references: GB:D13663; NID:g287439; PIDN:BAA02831.1; PID:g303895

C;Genetics:

A;Genome: plasmid

C;Superfamily: Shigella flexneri spa29 protein

C;Keywords: transmembrane protein

Query Match 62.9%; Score 39; DB 2; Length 256;  
Best Local Similarity 77.8%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 9  
||||||:|  
Db 27 LFFFLPFLN 35

RESULT 8

T21247

hypothetical protein F22B8.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C;Accession: T21247

R;McMurray, A.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z19396

A;Accession: T21247

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-359 <WIL>

A;Cross-references: EMBL:Z83106; PIDN:CAB05493.1; GSPDB:GN00023; CESP:F22B8.1

A;Experimental source: clone F22B8

C;Genetics:



A;Gene: CESP:F22B8.1  
A;Map position: 5  
A;Introns: 96/3; 132/2; 161/3; 200/3

Query Match 62.9%; Score 39; DB 2; Length 359;  
Best Local Similarity 70.0%; Pred. No. 27;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNV 10  
| | | | | : |  
Db 312 LFFFLPIFGV 321

RESULT 9

S74688

hypothetical protein sl11200 - *Synechocystis* sp. (strain PCC 6803)

C;Species: *Synechocystis* sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C;Accession: S74688

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.;  
Miyajima, N.; Hirose, M.; Sugiura, M.; Sasamoto, S.; Kimura, T.; Hosouchi, T.;  
Matsuno, A.; Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimpo, S.;  
Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium  
*Synechocystis* sp. PCC6803. II. Sequence determination of the entire genome and  
assignment of potential protein-coding regions.

A;Reference number: S74322; MUID:97061201; PMID:8905231

A;Accession: S74688

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-391 <KAN>

A;Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BAA16839.1;  
PID:d1017572; PID:g1651913

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June  
1996

Query Match 62.9%; Score 39; DB 2; Length 391;  
Best Local Similarity 70.0%; Pred. No. 30;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNV 10  
| | | | | : | |  
Db 99 LVFFFLPIANV 108

RESULT 10

F72257

lipopolysaccharide biosynthesis protein-related protein - *Thermotoga maritima*  
(strain MSB8)

C;Species: *Thermotoga maritima*

C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C;Accession: F72257

R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft,  
D.H.; Hickey, E.K.; Peterson, J.D.; Nelson, W.C.; Ketchum, K.A.; McDonald, L.;  
Utterback, T.R.; Malek, J.A.; Linher, K.D.; Garrett, M.M.; Stewart, A.M.;

Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; Heidelberg, J.; Sutton, G.G.; Fleischmann, R.D.; White, O.; Salzberg, S.L.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Nature 399, 323-329, 1999

A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*.

A;Reference number: A72200; MUID:99287316; PMID:10360571

A;Accession: F72257

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-471 <ARN>

A;Cross-references: GB:AE001793; GB:AE000512; NID:g4981963; PIDN:AAD36476.1; PID:g4981969; TIGR:TM1405

A;Experimental source: strain MSB8

C;Genetics:

A;Gene: TM1405

Query Match 62.9%; Score 39; DB 2; Length 471;  
Best Local Similarity 100.0%; Pred. No. 36;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 FLPVVNVL 11  
|||||||  
Db 153 FLPVVNVL 160

#### RESULT 11

E83002

drug efflux transporter PA5160 [imported] - *Pseudomonas aeruginosa* (strain PAO1)

C;Species: *Pseudomonas aeruginosa*

C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 03-Jun-2002

C;Accession: E83002

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber, R.L.; Goltry, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.; Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.; Hancock, R.E.W.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen.

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: E83002

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-509 <STO>

A;Cross-references: GB:AE004928; GB:AE004091; NID:g9951450; PIDN:AAG08545.1; GSPDB:GN00131; PASP:PA5160

A;Experimental source: strain PAO1

C;Genetics:

A;Gene: PA5160

C;Superfamily: lincomycin-resistance protein lmrB

Query Match 62.9%; Score 39; DB 2; Length 509;  
Best Local Similarity 50.0%; Pred. No. 39;  
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy            2 FFFLPVVNVL 11  
              |||:|:::|  
Db            376 FFFMPILSIL 385

RESULT 12

G31277

quate transport protein - *Neurospora crassa* (tentative sequence)

N;Alternate names: quinate transporter

C;Species: *Neurospora crassa*

C;Date: 26-Apr-1989 #sequence\_revision 26-Apr-1989 #text\_change 31-Mar-2000

C;Accession: S04254; G31277

R;Geever, R.F.; Huiet, L.; Baum, J.A.; Tyler, B.M.; Patel, V.B.; Rutledge, B.J.;  
Case, M.E.; Giles, N.H.

J. Mol. Biol. 207, 15-34, 1989

A;Title: DNA sequence, organization and regulation of the qa gene cluster of  
*Neurospora crassa*.

A;Reference number: S04250; MUID:89293848; PMID:2525625

A;Accession: S04254

A;Molecule type: DNA

A;Residues: 1-537 <GE2>

A;Cross-references: EMBL:X14603; NID:g3060; PIDN:CAA32752.1; PID:g3065

C;Genetics:

A;Gene: qa-y

C;Superfamily: maltose transport protein MAL61

C;Keywords: transmembrane protein

F;22-42/Domain: transmembrane #status predicted <TM01>

F;67-87/Domain: transmembrane #status predicted <TM02>

F;99-119/Domain: transmembrane #status predicted <TM03>

F;132-152/Domain: transmembrane #status predicted <TM04>

F;161-181/Domain: transmembrane #status predicted <TM05>

F;195-215/Domain: transmembrane #status predicted <TM06>

F;286-306/Domain: transmembrane #status predicted <TM07>

F;324-344/Domain: transmembrane #status predicted <TM08>

F;356-376/Domain: transmembrane #status predicted <TM09>

F;390-410/Domain: transmembrane #status predicted <TM10>

F;427-447/Domain: transmembrane #status predicted <TM11>

F;459-479/Domain: transmembrane #status predicted <TM12>

Query Match                    62.9%;    Score 39;    DB 2;    Length 537;

Best Local Similarity    50.0%;    Pred. No. 41;

Matches        6;    Conservative        3;    Mismatches        3;    Indels        0;    Gaps        0;

Qy            1 LFFFLPVVNVL 12  
              ::|||||    :|  
Db            475 IYFFLPVTKSIP 486

RESULT 13

T52573

cyclic nucleotide and calmodulin-regulated ion channel [imported] - *Arabidopsis thaliana*

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 24-Oct-2000

C;Accession: T52573

R;Kohler, C.; Merkle, T.; Neuhaus, G.

Plant J. 18, 97-104, 1999

A;Title: Characterisation of a novel gene family of putative cyclic nucleotide- and calmodulin-regulated ion channels in *Arabidopsis thaliana*.  
A;Reference number: Z26120  
A;Accession: T52573  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-710 <KOH>  
A;Cross-references: EMBL:Y17913; PIDN:CAB40130.1  
A;Experimental source: cultivar Columbia  
C;Genetics:  
A;Gene: cngc5

Query Match 62.9%; Score 39; DB 2; Length 710;  
Best Local Similarity 75.0%; Pred. No. 54;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFFLPVWN 9  
||:||||:  
Db 110 FFYLPVIN 117

#### RESULT 14

AD2302

hypothetical protein all3971 [imported] - *Nostoc* sp. (strain PCC 7120)

C;Species: *Nostoc* sp. PCC 7120

A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C;Accession: AD2302

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120.

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AD2302

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-364 <KUR>

A;Cross-references: GB:BA000019; PIDN:BAB75670.1; PID:g17133105; GSPDB:GN00179

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: all3971

Query Match 62.1%; Score 38.5; DB 2; Length 364;  
Best Local Similarity 47.4%; Pred. No. 34;  
Matches 9; Conservative 2; Mismatches 1; Indels 7; Gaps 1;

Qy 1 LFFF-----LPVNVLP 12  
|||| | |::||  
Db 281 LFFFAALISINLAVINILP 299

#### RESULT 15

AF3483

heme exporter protein B [imported] - *Brucella melitensis* (strain 16M)

C;Species: *Brucella melitensis*  
C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002  
C;Accession: AF3483  
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.;  
Ivanova, N.; Anderson, I.; Bhattacharyya, A.; Lykidis, A.; Reznik, G.;  
Jablonski, L.; Larsen, N.; D'Souza, M.; Bernal, A.; Mazur, M.; Goltsman, E.;  
Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesson, J.J.; Haselkorn,  
R.; Kyrpides, N.; Overbeek, R.  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A;Title: The genome sequence of the facultative intracellular pathogen *Brucella*  
*melitensis*.  
A;Reference number: AD3252; PMID:11756688  
A;Accession: AF3483  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-221 <KUR>  
A;Cross-references: GB:AE008917; PIDN:AAL53033.1; PID:g17983891; GSPDB:GN00190  
A;Experimental source: strain 16M  
C;Genetics:  
A;Gene: BMEI1852  
A;Map position: I  
C;Superfamily: cytochrome c biogenesis protein CycW

Query Match 61.3%; Score 38; DB 2; Length 221;  
Best Local Similarity 50.0%; Pred. No. 25;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
: ||| |::|:|  
Db 23 ILFFLAVISVMP 34

Search completed: August 24, 2004, 15:52:47  
Job time : 14.6418 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:51:19 ; Search time 43.4328 Seconds  
(without alignments)  
86.825 Million cell updates/sec

Title: US-09-641-801-4  
Perfect score: 62  
Sequence: 1 LFFFLPVVNVLP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 314255058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

8

Result	Query					
No.	Score	Match	Length	DB	ID	Description

1	62	100.0	12	14	US-10-281-652-4	Sequence 4, Appli
2	45	72.6	96	12	US-10-424-599-261912	Sequence 261912,
3	44	71.0	47	12	US-10-424-599-211893	Sequence 211893,
4	43	69.4	54	12	US-10-424-599-143141	Sequence 143141,
5	43	69.4	180	12	US-10-424-599-227817	Sequence 227817,
6	43	69.4	197	12	US-10-424-599-218474	Sequence 218474,
7	43	69.4	678	12	US-10-424-599-260138	Sequence 260138,
8	42	67.7	72	12	US-10-424-599-197974	Sequence 197974,
9	42	67.7	101	12	US-10-424-599-245079	Sequence 245079,
10	41	66.1	54	12	US-10-424-599-234012	Sequence 234012,
11	41	66.1	56	12	US-10-424-599-154599	Sequence 154599,
12	41	66.1	81	12	US-10-424-599-228846	Sequence 228846,
13	41	66.1	82	12	US-10-424-599-252118	Sequence 252118,
14	41	66.1	126	15	US-10-094-749-1945	Sequence 1945, Ap
15	41	66.1	141	12	US-10-424-599-272793	Sequence 272793,
16	41	66.1	153	12	US-10-424-599-171577	Sequence 171577,
17	41	66.1	666	16	US-10-437-963-138531	Sequence 138531,
18	41	66.1	738	16	US-10-437-963-138533	Sequence 138533,
19	40	64.5	98	12	US-10-425-114-42601	Sequence 42601, A
20	40	64.5	132	16	US-10-437-963-187827	Sequence 187827,
21	40	64.5	363	16	US-10-437-963-167680	Sequence 167680,
22	39	62.9	46	12	US-10-424-599-149052	Sequence 149052,
23	39	62.9	53	12	US-10-424-599-208834	Sequence 208834,
24	39	62.9	58	12	US-10-424-599-181046	Sequence 181046,
25	39	62.9	95	12	US-10-424-599-237727	Sequence 237727,
26	39	62.9	118	12	US-10-424-599-164024	Sequence 164024,
27	39	62.9	356	16	US-10-437-963-167720	Sequence 167720,
28	39	62.9	422	12	US-10-282-122A-49461	Sequence 49461, A
29	38	61.3	68	12	US-10-424-599-204168	Sequence 204168,
30	38	61.3	68	16	US-10-437-963-140548	Sequence 140548,
31	38	61.3	85	12	US-10-424-599-244849	Sequence 244849,
32	38	61.3	108	16	US-10-437-963-192194	Sequence 192194,
33	38	61.3	563	16	US-10-437-963-107198	Sequence 107198,
34	37	59.7	42	10	US-09-933-767-456	Sequence 456, App
35	37	59.7	42	12	US-10-004-860-456	Sequence 456, App
36	37	59.7	42	14	US-10-023-282-456	Sequence 456, App
37	37	59.7	52	16	US-10-437-963-199617	Sequence 199617,
38	37	59.7	53	12	US-10-424-599-235499	Sequence 235499,
39	37	59.7	63	14	US-10-083-357-932	Sequence 932, App
40	37	59.7	72	12	US-10-424-599-256914	Sequence 256914,
41	37	59.7	84	9	US-09-764-877-1107	Sequence 1107, Ap
42	37	59.7	84	15	US-10-242-515-1107	Sequence 1107, Ap
43	37	59.7	89	16	US-10-437-963-109671	Sequence 109671,
44	37	59.7	102	16	US-10-437-963-142097	Sequence 142097,
45	37	59.7	103	10	US-09-933-767-251	Sequence 251, App

#### ALIGNMENTS

RESULT 1  
 US-10-281-652-4  
 ; Sequence 4, Application US/10281652  
 ; Publication No. US20030091606A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: STANTON, G. John

```

; APPLICANT: HUGHES, Thomas K.
; APPLICANT: BOLDOGH, Istvan
; TITLE OF INVENTION: USE OF COLOSTRININ, CONSTITUENT PEPTIDES THEREOF, AND
; TITLE OF INVENTION: ANALOGS THEREOF AS OXIDATIVE STRESS REGULATORS
; FILE REFERENCE: 265.00220101
; CURRENT APPLICATION NUMBER: US/10/281,652
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: US/09/641,803
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/149,310
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-10-281-652-4

```

```

Query Match          100.0%; Score 62; DB 14; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 LFFFLPVVNVLP 12
        |||||
Db      1 LFFFLPVVNVLP 12

```

## RESULT 2

```

US-10-424-599-261912
; Sequence 261912, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 261912
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_78529C.1.pep
US-10-424-599-261912

```

```

Query Match          72.6%; Score 45; DB 12; Length 96;
Best Local Similarity 72.7%; Pred. No. 5.1;

```



Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12  
||| |||:| |  
Db 12 FFFFPVVSVEP 22

RESULT 3

US-10-424-599-211893  
; Sequence 211893, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated  
With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 211893  
; LENGTH: 47  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_33366C.1.pep  
US-10-424-599-211893

Query Match 71.0%; Score 44; DB 12; Length 47;  
Best Local Similarity 66.7%; Pred. No. 3.7;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVNVLP 12  
||||| :|:|  
Db 34 LFFFLDKINLLP 45

RESULT 4

US-10-424-599-143141  
; Sequence 143141, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated  
With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 143141

; LENGTH: 54  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_10026C.1.pep  
US-10-424-599-143141

Query Match 69.4%; Score 43; DB 12; Length 54;  
Best Local Similarity 63.6%; Pred. No. 6.2;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12  
|| ||: | ||  
Db 39 FFILPITNALP 49

RESULT 5

US-10-424-599-227817  
; Sequence 227817, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated  
With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 227817  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_47747C.1.pep  
US-10-424-599-227817

Query Match 69.4%; Score 43; DB 12; Length 180;  
Best Local Similarity 77.8%; Pred. No. 21;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 9  
|||:|||:|  
Db 125 LFFYLPVIN 133

RESULT 6

US-10-424-599-218474  
; Sequence 218474, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua

```
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 218474
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(197)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_3930C.1.pep
US-10-424-599-218474
```

```
Query Match          69.4%; Score 43; DB 12; Length 197;
Best Local Similarity 77.8%; Pred. No. 23;
Matches      7; Conservative 2; Mismatches      0; Indels      0; Gaps      0;
```

```
Qy      1 LFFFLPVVN 9
        |||:|||:|
Db      126 LFFYLPVIN 134
```

#### RESULT 7

```
US-10-424-599-260138
; Sequence 260138, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 260138
; LENGTH: 678
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(678)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_76929C.1.pep
US-10-424-599-260138
```

Query Match 69.4%; Score 43; DB 12; Length 678;  
Best Local Similarity 77.8%; Pred. No. 79;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 9  
|||:|:|:  
Db 74 LFFYLPVIN 82

RESULT 8

US-10-424-599-197974  
; Sequence 197974, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated  
With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 197974  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_20796C.1.pep  
US-10-424-599-197974

Query Match 67.7%; Score 42; DB 12; Length 72;  
Best Local Similarity 63.6%; Pred. No. 12;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12  
||||::| ||  
Db 15 FFFLEIINSLP 25

RESULT 9

US-10-424-599-245079  
; Sequence 245079, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated  
With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 245079  
; LENGTH: 101  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_63337C.1.pep  
US-10-424-599-245079

Query Match 67.7%; Score 42; DB 12; Length 101;  
Best Local Similarity 87.5%; Pred. No. 17;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFFLPVVN 9  
||||:||  
Db 76 FFFLPIVN 83

RESULT 10

US-10-424-599-234012  
; Sequence 234012, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated  
With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 234012  
; LENGTH: 54  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_53337C.1.pep  
US-10-424-599-234012

Query Match 66.1%; Score 41; DB 12; Length 54;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVV 8  
|||||||  
Db 8 LFFFLPVV 15

RESULT 11

US-10-424-599-154599  
; Sequence 154599, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J

```
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154599
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110624C.1.pep
US-10-424-599-154599
```

```
Query Match          66.1%;  Score 41;  DB 12;  Length 56;
Best Local Similarity 63.6%;  Pred. No. 14;
Matches      7;  Conservative    2;  Mismatches    2;  Indels      0;  Gaps      0;
```

```
Qy      1 LFFFLPVVNVL 11
        |||||:  :|
Db      18 LFFFLPISTIL 28
```

#### RESULT 12

US-10-424-599-228846

```
; Sequence 228846, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated
With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 228846
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48675C.1.pep
US-10-424-599-228846
```

```
Query Match          66.1%;  Score 41;  DB 12;  Length 81;
Best Local Similarity 71.4%;  Pred. No. 20;
Matches    10;  Conservative    0;  Mismatches    2;  Indels      2;  Gaps      1;
```

```
Qy      1 LFFFLPV--VNVLP 12
        | || | |||||
```

Db 54 LIFFFPVPHVNVLP 67

RESULT 13

US-10-424-599-252118

; Sequence 252118, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 252118

; LENGTH: 82

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_69690C.1.pep

US-10-424-599-252118

Query Match 66.1%; Score 41; DB 12; Length 82;

Best Local Similarity 60.0%; Pred. No. 20;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 FFLPVVNVLP 12

| :|:|:|

Db 11 FLVPIVNILP 20

RESULT 14

US-10-094-749-1945

; Sequence 1945, Application US/10094749

; Publication No. US20030219741A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, RYOTARO

; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAOHIKO

; APPLICANT: YOSHIKAWA, TSUTOMU

; APPLICANT: OTSUKA, MOTUYUKI

; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1945  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-094-749-1945

Query Match 66.1%; Score 41; DB 15; Length 126;  
Best Local Similarity 72.7%; Pred. No. 31;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12  
||||| |: ||  
Db 35 FFFLPVSSLP 45

RESULT 15

US-10-424-599-272793

; Sequence 272793, Application US/10424599  
; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 272793

; LENGTH: 141  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_88354C.1.pep  
US-10-424-599-272793

Query Match 66.1%; Score 41; DB 12; Length 141;  
Best Local Similarity 63.6%; Pred. No. 35;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12  
||| |: |||



Db

37 FFFFPISQVLP 47

Search completed: August 24, 2004, 16:41:18  
Job time : 45.4328 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 15:23:00 ; Search time 37.0746 Seconds  
(without alignments)  
102.124 Million cell updates/sec

Title: US-09-641-801-4  
Perfect score: 62  
Sequence: 1 LFFFLPVVNVLP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriaphage:\*  
17: sp\_archaeophages:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	
No.	Score Match Length DB ID	Description

---

1	48	77.4	469	16	Q8YDU8	Q8ydu8 brucella me
2	48	77.4	469	16	Q8FXP8	Q8fxp8 brucella su
3	45	72.6	144	10	Q98S86	Q98s86 guillardia
4	44	71.0	62	16	Q8UHQ7	Q8uhq7 agrobacteri
5	44	71.0	218	8	Q9G873	Q9g873 malawimonas
6	44	71.0	249	10	Q24102	Q24102 medicago tr
7	44	71.0	271	16	Q8FUV9	Q8fuv9 brucella su
8	44	71.0	288	2	Q8VQL7	Q8vql7 brucella ab
9	44	71.0	525	17	Q9HIU9	Q9hiu9 thermoplasm
10	42	67.7	545	16	Q8A411	Q8a411 bacteroides
11	41	66.1	66	8	Q36231	Q36231 williopsis
12	41	66.1	126	4	Q96NK3	Q96nk3 homo sapien
13	41	66.1	376	5	Q8IM69	Q8im69 plasmodium
14	41	66.1	488	16	Q67391	Q67391 aquifex ae
15	41	66.1	692	3	Q06665	Q06665 saccharomyc
16	40	64.5	288	10	Q7XJX5	Q7xjx5 oryza sativ
17	40	64.5	352	17	Q8PUB3	Q8pub3 methanosarc
18	40	64.5	431	16	Q898L1	Q898l1 clostridium
19	40	64.5	458	5	Q9XWU8	Q9xwu8 caenorhabdi
20	40	64.5	471	2	Q9EYG6	Q9eyg6 actinobacil
21	40	64.5	724	10	Q7XTN4	Q7xtn4 oryza sativ
22	40	64.5	1441	16	Q89TS8	Q89ts8 bradyrhizob
23	39	62.9	62	16	Q819H8	Q819h8 bacillus ce
24	39	62.9	104	17	Q8PVK3	Q8pvk3 methanosarc
25	39	62.9	305	16	Q88GD7	Q88gd7 pseudomonas
26	39	62.9	359	5	O17830	O17830 caenorhabdi
27	39	62.9	391	16	P72824	P72824 synechocyst
28	39	62.9	430	2	Q8VNV9	Q8vvn9 clostridium
29	39	62.9	430	16	Q8XHA3	Q8xha3 clostridium
30	39	62.9	431	5	Q9BMP8	Q9bmp8 plasmodium
31	39	62.9	436	2	Q8KR72	Q8kr72 photorhabdu
32	39	62.9	471	16	Q9X1C5	Q9xlc5 thermotoga
33	39	62.9	509	16	O85163	O85163 pseudomonas
34	39	62.9	552	16	Q9CPI8	Q9cpi8 pasteurella
35	39	62.9	710	10	Q9XFS3	Q9xfs3 arabidopsis
36	39	62.9	717	10	Q8RWS9	Q8rws9 arabidopsis
37	39	62.9	4524	5	Q8I3J9	Q8i3j9 plasmodium
38	38.5	62.1	177	10	Q7XNL6	Q7xnl6 oryza sativ
39	38.5	62.1	363	2	Q8GMR6	Q8gmr6 synechococc
40	38	61.3	108	10	Q84Q99	Q84q99 oryza sativ
41	38	61.3	215	8	Q9TDV7	Q9tdv7 astreopora
42	38	61.3	221	16	Q8YEM6	Q8yem6 brucella me
43	38	61.3	221	16	Q8G357	Q8g357 brucella su
44	38	61.3	296	16	Q8UBB9	Q8ubb9 agrobacteri
45	38	61.3	351	17	Q8TQG7	Q8tqg7 methanosarc

# ALIGNMENTS

## RESULT 1

Q8YDU8

ID Q8YDU8 PRELIMINARY; PRT; 469 AA.  
AC Q8YDU8;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Enterobactin synthetase component F.  
 GN BMEII0076.  
 OS Brucella melitensis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29459;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=16M / ATCC 23456 / Biotype 1;  
 RX MEDLINE=20020109; PubMed=11756688;  
 RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,  
 RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,  
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,  
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
 RA Haselkorn R., Kyrpides N., Overbeek R.;  
 RT "The genome sequence of the facultative intracellular pathogen  
 RT Brucella melitensis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).  
 DR EMBL; AE009646; AAL53317.1; -.  
 DR PIR; AB3519; AB3519.  
 DR InterPro; IPR001242; Condensatn.  
 DR Pfam; PF00668; Condensation; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 469 AA; 53074 MW; 17A7B73A02428D46 CRC64;

Query Match 77.4%; Score 48; DB 16; Length 469;  
 Best Local Similarity 72.7%; Pred. No. 3.4;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12  
 ||| |::|||  
 Db 374 FFFSPLINVLP 384

## RESULT 2

### Q8FXP8

ID Q8FXP8 PRELIMINARY; PRT; 469 AA.  
 AC Q8FXP8;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Enterobactin synthetase, component F, putative.  
 GN BRA0017.  
 OS Brucella suis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29461;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1330 / Biovar 1;  
 RX MEDLINE=22247741; PubMed=12271122;  
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,  
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,  
 RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,  
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,  
 RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,  
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;

RT "The Brucella suis genome reveals fundamental similarities between  
RT animal and plant pathogens and symbionts."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).  
DR EMBL; AE014506; AAN33229.1; -.  
DR TIGR; BRA0017; -.  
DR InterPro; IPR001242; Condensatn.  
DR Pfam; PF00668; Condensation; 1.  
KW Complete proteome.  
SQ SEQUENCE 469 AA; 52986 MW; AD46038DCF854A31 CRC64;

Query Match 77.4%; Score 48; DB 16; Length 469;  
Best Local Similarity 72.7%; Pred. No. 3.4;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVLP 12  
||| |::|||  
Db 374 FFFSPLINVLP 384

#### RESULT 3

Q98S86

ID Q98S86 PRELIMINARY; PRT; 144 AA.  
AC Q98S86;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein orf144 from chromosome 3.  
GN ORF144.  
OS Guillardia theta (Cryptomonas phi).  
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
OX NCBI\_TaxID=55529;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21223349; PubMed=11323671;  
RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,  
RA Wu X., Reith M., Cavalier-Smith T., Maier U.G.;  
RT "The highly reduced genome of an enslaved algal nucleus."  
RL Nature 410:1091-1096(2001).  
DR EMBL; AF083031; AAK39696.1; -.  
DR PIR; D90125; D90125.  
KW Hypothetical protein.  
SQ SEQUENCE 144 AA; 17625 MW; 72649208661F4DA5 CRC64;

Query Match 72.6%; Score 45; DB 10; Length 144;  
Best Local Similarity 66.7%; Pred. No. 3.9;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVNVLP 12  
:|||| ||:||  
Db 24 IFFFLKKVNILP 35

#### RESULT 4

Q8UHQ7

ID Q8UHQ7 PRELIMINARY; PRT; 62 AA.  
AC Q8UHQ7;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein Atu0623.  
 GN ATU0623.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 OX NCBI\_TaxID=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kuttyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 RT C58.";  
 RL Science 294:2317-2323(2001).  
 DR EMBL; AE009030; AAL41639.1; -.  
 DR PIR; AI2652; AI2652.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 62 AA; 7135 MW; C67C9F18234FEAB6 CRC64;

Query Match 71.0%; Score 44; DB 16; Length 62;  
 Best Local Similarity 72.7%; Pred. No. 2.8;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 FFFLPVVNVLP 12  
 | : | | | : | | | |  
 Db 5 FYSLPVMNVLP 15

# RESULT 5

Q9G873

ID Q9G873 PRELIMINARY; PRT; 218 AA.  
 AC Q9G873;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE ABC transporter channel subunit.  
 GN YEJV.  
 OS Malawimonas jakobiformis.  
 OG Mitochondrion.  
 OC Eukaryota; Malawimonadidae; Malawimonas.  
 OX NCBI\_TaxID=136089;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Burger G., O'Kelly C.J., Gray M.W., Lang B.F.;  
 RT "Comparative analysis of mitochondrial genomes of the ancient jakobid  
 RT protists.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AF295546; AAG13700.1; -.  
DR GO; GO:0005739; C:mitochondrion; IEA.  
KW Mitochondrion.  
SQ SEQUENCE 218 AA; 25870 MW; EF7A162FEF88D674 CRC64;

Query Match 71.0%; Score 44; DB 8; Length 218;  
Best Local Similarity 54.5%; Pred. No. 8.5;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVL 11  
|||:::  
Db 101 LFFFLPIITII 111

#### RESULT 6

O24102

ID O24102 PRELIMINARY; PRT; 249 AA.  
AC O24102;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE MtN4 protein (Fragment).  
GN MTN4.  
OS Medicago truncatula (Barrel medic).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
OX NCBI\_TaxID=3880;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Jemalong J5; TISSUE=Root nodules;  
RA Gamas P.;  
RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Jemalong J5; TISSUE=Root nodules;  
RX MEDLINE=96212994; PubMed=8634476;  
RA Gamas P., de Carvalho Niebel F., Lescure N., Cullimore J.;  
RT "Use of a subtractive hybridization approach to identify new Medicago  
truncatula genes induced during root nodule development."  
RL Mol. Plant Microbe Interact. 9:233-242(1996).  
DR EMBL; Y15372; CAA75594.1; -.  
DR HSSP; P24337; 1HYP.  
DR InterPro; IPR003612; AAI.  
DR Pfam; PF00234; tryp\_alpha\_amyl; 1.  
DR SMART; SM00499; AAI; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 249 AA; 26923 MW; 4BF9256A0FDD1318 CRC64;

Query Match 71.0%; Score 44; DB 10; Length 249;  
Best Local Similarity 66.7%; Pred. No. 9.6;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVL 12  
:||:||||  
Db 107 MFFYLPVVPVTP 118

# RESULT 7

Q8FUV9

ID Q8FUV9 PRELIMINARY; PRT; 271 AA.  
AC Q8FUV9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Polyamine ABC transporter, permease protein, putative.  
GN BRA1106.  
OS Brucella suis.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Brucellaceae; Brucella.  
OX NCBI\_TaxID=29461;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1330 / Biovar 1;  
RX MEDLINE=22247741; PubMed=12271122;  
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,  
RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,  
RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,  
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,  
RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,  
RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;  
RT "The Brucella suis genome reveals fundamental similarities between  
RT animal and plant pathogens and symbionts.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).  
DR EMBL; AE014602; AAN34268.1; -.  
DR TIGR; BRA1106; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000515; BPD\_transp.  
DR Pfam; PF00528; BPD\_transp; 1.  
DR PROSITE; PS00402; BPD\_TRANSP\_INN\_MEMBR; 1.  
KW Complete proteome.  
SQ SEQUENCE 271 AA; 29759 MW; 6DDEC475BD4C1A83 CRC64;

Query Match 71.0%; Score 44; DB 16; Length 271;  
Best Local Similarity 72.7%; Pred. No. 10;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVL 11  
| | | | | : | | : |  
Db 14 LFFFYPLVNLL 24

# RESULT 8

Q8VQL7

ID Q8VQL7 PRELIMINARY; PRT; 288 AA.  
AC Q8VQL7;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Putative ABC transporter permease protein B.  
GN BATN1953.ORF9.  
OS Brucella abortus.



OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=235;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=544;  
 RA Bricker B.J.;  
 RT "Tn1953, a new element from Brucella abortus.";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM.  
 CC PROBABLY RESPONSIBLE FOR THE TRANSLOCATION OF THE SUBSTRATE ACROSS  
 CC THE MEMBRANE (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT  
 CC SYSTEM PERMEASE FAMILY.  
 DR EMBL; AF454951; AAL59331.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000515; BPD\_transp.  
 DR Pfam; PF00528; BPD\_transp; 1.  
 DR PROSITE; PS00402; BPD\_TRANSP\_INN\_MEMBR; 1.  
 KW Transmembrane; Transport.  
 SQ SEQUENCE 288 AA; 31696 MW; B5C20EA208DCFD8E CRC64;

Query Match 71.0%; Score 44; DB 2; Length 288;  
 Best Local Similarity 72.7%; Pred: No. 11;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVL 11  
 |||| |:|:|  
 Db 31 LFFFYPLVNLL 41

#### RESULT 9

##### Q9HIU9

ID Q9HIU9 PRELIMINARY; PRT; 525 AA.  
 AC Q9HIU9;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Cytochrome b related protein.  
 GN CYTB OR TA1228.  
 OS Thermoplasma acidophilum.  
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;  
 OC Thermoplasmataceae; Thermoplasma.  
 OX NCBI\_TaxID=2303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 1728;  
 RX MEDLINE=20479972; PubMed=11029001;  
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,  
 RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;  
 RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma  
 RT acidophilum.";  
 RL Nature 407:508-513(2000).  
 CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE

CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A  
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL  
 CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).  
 CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
 CC BOUND TO THE PROTEIN (BY SIMILARITY).  
 CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,  
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.  
 DR EMBL; AL445066; CAC12352.1; -.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR005797; Cytb\_b6\_N.  
 DR Pfam; PF00033; cytochrome\_b\_N; 1.  
 KW Electron transport; Heme; Respiratory chain; Transmembrane;  
 KW Complete proteome.  
 SQ SEQUENCE 525 AA; 58544 MW; 145564FA78C665B7 CRC64;

Query Match 71.0%; Score 44; DB 17; Length 525;  
 Best Local Similarity 75.0%; Pred. No. 19;  
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
 ||| ||:| |||  
 Db 368 LFFILPLVIVLP 379

# RESULT 10

Q8A411

ID Q8A411 PRELIMINARY; PRT; 545 AA.  
 AC Q8A411;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative MFS transporter.  
 GN BT2793.  
 OS Bacteroides thetaiotaomicron.  
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 OC Bacteroidaceae; Bacteroides.  
 OX NCBI\_TaxID=818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VPI-5482 / ATCC 29148;  
 RX MEDLINE=22550858; PubMed=12663928;  
 RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
 RA Chiang H.C., Hooper L.V., Gordon J.I.;  
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";  
 RL Science 299:2074-2076(2003).  
 DR EMBL; AE016937; AAO77899.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000412; ABC\_transpt2.  
 DR PROSITE; PS00890; ABC2\_MEMBRANE; 1.  
 KW Complete proteome.

SQ SEQUENCE 545 AA; 61484 MW; 976E2EF859984466 CRC64;

Query Match 67.7%; Score 42; DB 16; Length 545;  
Best Local Similarity 70.0%; Pred. No. 43;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNV 10  
|||||:::|  
Db 24 LFFFLPILSV 33

RESULT 11

Q36231

ID Q36231 PRELIMINARY; PRT; 66 AA.  
AC Q36231;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Partial putative ORF (Fragment).  
OS Williopsis saturnus var. suaveolens.  
OG Mitochondrion.  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Williopsis.  
OX NCBI\_TaxID=58637;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CBS 255;  
RA Nosek J.;  
RL Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; X77238; CAA54455.1; -.  
DR PIR; S49114; S49114.  
DR GO; GO:0005739; C:mitochondrion; IEA.  
KW Mitochondrion.  
FT NON\_TER 1 1  
SQ SEQUENCE 66 AA; 7586 MW; 78C9BCF9A31B94FB CRC64;

Query Match 66.1%; Score 41; DB 8; Length 66;  
Best Local Similarity 50.0%; Pred. No. 9.8;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12  
||||: :: |:|  
Db 37 LFFFIMIIGVMP 48

RESULT 12

Q96NK3

ID Q96NK3 PRELIMINARY; PRT; 126 AA.  
AC Q96NK3;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein FLJ30690.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,  
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,  
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,  
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,  
 RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,  
 RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,  
 RA Masuho Y., Nagai K., Isogai T.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AK055252; BAB70890.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 126 AA; 14391 MW; D1B23CDB0644847B CRC64;

Query Match 66.1%; Score 41; DB 4; Length 126;  
 Best Local Similarity 72.7%; Pred. No. 17;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FFFFLPVVNVLP 12  
 ||||| |: ||  
 Db 35 FFFFLPPVSSLP 45

# RESULT 13

Q8IM69

ID Q8IM69 PRELIMINARY; PRT; 376 AA.  
 AC Q8IM69;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Exopolyphosphatase, putative.  
 GN PF14\_0022.  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=3D7;  
 RX MEDLINE=22255705; PubMed=12368864;  
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,  
 RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,  
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,  
 RA Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,  
 RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,  
 RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,  
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,  
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,  
 RA Fraser C.M., Barrell B.;  
 RT "Genome sequence of the human malaria parasite Plasmodium  
 RT falciparum.";  
 RL Nature 419:498-511(2002).  
 DR EMBL; AE014816; AAN36634.1; -.  
 SQ SEQUENCE 376 AA; 45051 MW; 0FDE636678149F2C CRC64;

Query Match 66.1%; Score 41; DB 5; Length 376;  
Best Local Similarity 60.0%; Pred. No. 46;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNV 10  
| ||:||:|:  
Db 158 LMFFIPVINI 167

RESULT 14

O67391

ID O67391 PRELIMINARY; PRT; 488 AA.  
AC O67391;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE NADH dehydrogenase I chain N.  
GN NUON2 OR AQ\_1383.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.  
OX NCBI\_TaxID=63363;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VF5;  
RX MEDLINE=98196666; PubMed=9537320;  
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,  
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
RT "The complete genome of the hyperthermophilic bacterium Aquifex  
RT aeolicus.";  
RL Nature 392:353-358(1998).  
CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: TO NADH-UBIQUINONE/PLASTOQUINONE (COMPLEX I), VARIOUS  
CC CHAINS.  
CC -!- SIMILARITY: TO ONE OF THE POLYPEPTIDE CHAINS OF THE NADH-UBIQUINOL  
CC OXIDOREDUCTASE OF CHLOROPLASTS OR MITOCHONDRIA.  
DR EMBL; AE000737; AAC07354.1; -.  
DR PIR; D70420; D70420.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.  
DR InterPro; IPR003918; NADHub\_oxred4.  
DR InterPro; IPR003916; NADHub\_oxred5.  
DR InterPro; IPR001750; Oxidored\_q1.  
DR Pfam; PF00361; oxidored\_q1; 1.  
DR PRINTS; PR01434; NADHDHGNASE5.  
DR PRINTS; PR01437; NUOXDRDTASE4.  
KW NAD; Oxidoreductase; Transmembrane; Complete proteome.  
SQ SEQUENCE 488 AA; 54264 MW; 64EFE37967ED5435 CRC64;

Query Match 66.1%; Score 41; DB 16; Length 488;  
Best Local Similarity 58.3%; Pred. No. 58;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12

Db                    | ||:|:| | :|  
255 LAFFIPLVRVMP 266

RESULT 15

Q06665

ID    Q06665            PRELIMINARY;            PRT;    692 AA.  
AC    Q06665;  
DT    01-NOV-1996 (TrEMBLrel. 01, Created)  
DT    01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT    01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE    Chromosome IV COSMID 9740.  
GN    YDR314C OR D9740.21.  
OS    *Saccharomyces cerevisiae* (Baker's yeast).  
OC    Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC    Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX    NCBI\_TaxID=4932;  
RN    [1]  
RP    SEQUENCE FROM N.A.  
RA    Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,  
RA    Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,  
RA    Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,  
RA    Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,  
RA    Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,  
RA    Taich A., Trevaskis E., Vignati D., Wilcox L., Wohldman P., Vaudin M.,  
RA    Wilson R., Waterston R.;  
RL    Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.  
RN    [2]  
RP    SEQUENCE FROM N.A.  
RA    Ding H.;  
RL    Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.  
RN    [3]  
RP    SEQUENCE FROM N.A.  
RA    Waterston R.;  
RL    Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.  
RN    [4]  
RP    SEQUENCE FROM N.A.  
RA    Jia Y., Cherry J.M.;  
RL    Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.  
DR    EMBL; U28374; AAB64750.1; -.  
DR    PIR; S61200; S61200.  
DR    SGD; S0002722; YDR314C.  
DR    GO; GO:0005634; C:nucleus; IEA.  
DR    GO; GO:0003684; F:damaged DNA binding; IEA.  
DR    GO; GO:0006289; P:nucleotide-excision repair; IEA.  
DR    InterPro; IPR004583; Rad4.  
DR    Pfam; PF03835; Rad4; 1.  
SQ    SEQUENCE    692 AA;    82042 MW;    EF63BF1F002E7F2D CRC64;

Query Match                    66.1%;    Score 41;    DB 3;    Length 692;  
Best Local Similarity    66.7%;    Pred. No. 80;  
Matches        8;    Conservative        2;    Mismatches        2;    Indels        0;    Gaps        0;

Qy                    1 LFFFLPVVNVLP 12  
                      ||||: : ||||  
Db                    242 LFFFIILENVLP 253

Search completed: August 24, 2004, 15:50:36  
Job time : 42.0746 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 14:57:04 ; Search time 6.44776 Seconds  
(without alignments)  
96.908 Million cell updates/sec

Title: US-09-641-801-4  
Perfect score: 62  
Sequence: 1 LFFFLPVNVLP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		%	Query				
No.	Score	Match	Length	DB	ID	Description	
1	39	62.9	256	1	SPAR_SHIFL	P40706 shigella fl	
2	39	62.9	537	1	QAY_NEUCR	P11636 neurospora	
3	39	62.9	717	1	CNG5_ARATH	Q8rws9 arabidopsis	
4	38.5	62.1	364	1	Y3H1_ANASP	Q8yq64 anabaena sp	
5	38	61.3	461	1	YCJJ_ECOLI	P76037 escherichia	
6	38	61.3	678	1	CG15_ARATH	Q9sl29 arabidopsis	
7	38	61.3	738	1	CNG7_ARATH	Q9s9n5 arabidopsis	
8	38	61.3	753	1	CNG8_ARATH	Q9fxh6 arabidopsis	
9	37	59.7	263	1	SPAR_SALTY	P40701 salmonella	
10	37	59.7	514	1	PPCK_PSESM	Q88az4 pseudomonas	
11	37	59.7	546	1	LNT_TREPA	O83279 treponema p	
12	37	59.7	663	1	NKX1_CHICK	Q9ial8 gallus gall	
13	36	58.1	249	1	YD68_METJA	Q58763 methanococc	
14	36	58.1	302	1	MRAY_THEMA	Q9wy77 thermotoga	
15	36	58.1	304	1	YTR1_BUCSC	Q44601 buchnera ap	
16	36	58.1	306	1	CBPB_BOVIN	P00732 bos taurus	
17	36	58.1	327	1	CD1A_HUMAN	P06126 homo sapien	



18	36	58.1	404	1	O74A_DROME	Q9vvf3	drosophila
19	36	58.1	415	1	CBPB_RAT	P19223	rattus norv
20	36	58.1	485	1	YC11_KLEPN	Q48457	klebsiella
21	36	58.1	501	1	C723_ARATH	O65785	arabidopsis
22	36	58.1	696	1	CG13_ARATH	Q9ld40	arabidopsis
23	36	58.1	711	1	CG10_ARATH	Q9lnj0	arabidopsis
24	36	58.1	2555	1	FAFY_HUMAN	O00507	h probable
25	35	56.5	46	1	PSBK_CHLRE	P18263	chlamydomon
26	35	56.5	200	1	ATKC_YERPE	Q8zd98	yersinia pe
27	35	56.5	263	1	MURI_BUCBP	P59574	buchnera ap
28	35	56.5	270	1	T2C1_COREQ	P42827	corynebacte
29	35	56.5	305	1	PEX2_HUMAN	P28328	homo sapien
30	35	56.5	313	1	OXA2_HUMAN	Q8ngj7	homo sapien
31	35	56.5	313	1	OXA4_HUMAN	Q8ngj6	homo sapien
32	35	56.5	315	1	OXL1_HUMAN	Q8ngj5	homo sapien
33	35	56.5	348	1	OPSD_SARPU	P79902	sargocentro
34	35	56.5	351	1	OPSD_SARDI	P79898	sargocentro
35	35	56.5	352	1	OPSD_GOBNI	Q9ygz2	gobius nige
36	35	56.5	352	1	OPSD_POMMI	P35403	pomatoschis
37	35	56.5	352	1	OPSD_ZOSOP	Q9ygy9	zosterisess
38	35	56.5	366	1	GHSR_HUMAN	Q92847	homo sapien
39	35	56.5	382	1	ADH2_ECOLI	P37686	escherichia
40	35	56.5	410	1	NUOH_MYCTU	P95174	mycobacteri
41	35	56.5	429	1	RNE_GUITH	O78453	guillardia
42	35	56.5	436	1	Y326_METJA	Q57772	methanococc
43	35	56.5	462	1	SYSC_YEAST	P07284	saccharomyc
44	35	56.5	538	1	THIP_HAEIN	P44985	haemophilus
45	35	56.5	568	1	PTLB_LACLA	P23531	lactococcus

# ALIGNMENTS

## RESULT 1

### SPAR\_SHIFL

ID SPAR\_SHIFL STANDARD; PRT; 256 AA.  
AC P40706; Q8VSG7; Q9AFR9; Q9AJW0;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Surface presentation of antigens protein spar (Spa29 protein).  
GN SPAR OR SPA29 OR CP0155.  
OS Shigella flexneri, and  
OS Shigella sonnei.  
OG Plasmid pWR100, Plasmid pMYSH6000, and Plasmid pCP301.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Shigella.  
OX NCBI\_TaxID=623, 624;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=S.flexneri; STRAIN=M90T / Serotype 5; PLASMID=pWR100;  
RX MEDLINE=20566792; PubMed=11115111;  
RA Buchrieser C., Glaser P., Rusniok C., Nedjari H., d'Hauteville H.,  
RA Kunst F., Sansonetti P., Parsot C.;  
RT "The virulence plasmid pWR100 and the repertoire of proteins secreted  
RT by the type III secretion apparatus of Shigella flexneri.";  
RL Mol. Microbiol. 38:760-771(2000).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.flexneri; STRAIN=M90T / Serotype 5; PLASMID=pWR100;  
 RX MEDLINE=21189246; PubMed=11292750;  
 RA Venkatesan M.M., Goldberg M.B., Rose D.J., Grotbeck E.J., Burland V.,  
 RA Blattner F.R.;  
 RT "Complete DNA sequence and analysis of the large virulence plasmid of  
 RT Shigella flexneri.";  
 RL Infect. Immun. 69:3271-3285(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.flexneri; STRAIN=YSH6000 / Serotype 2a; PLASMID=pMYSH6000;  
 RX MEDLINE=93224456; PubMed=8385666;  
 RA Sasakawa C., Komatsu K., Tobe T., Suzuki T., Yoshikawa M.;  
 RT "Eight genes in region 5 that form an operon are essential for  
 RT invasion of epithelial cells by Shigella flexneri 2a.";  
 RL J. Bacteriol. 175:2334-2346(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a; PLASMID=pCP301;  
 RX MEDLINE=22272406; PubMed=12384590;  
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,  
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,  
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,  
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,  
 RA Yu J.;  
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity  
 RT through comparison with genomes of Escherichia coli K12 and O157.";  
 RL Nucleic Acids Res. 30:4432-4441(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=S.sonnei; STRAIN=HW383;  
 RA Arakawa E., Kato J.I., Ito K.I., Watanabe H.;  
 RT "Comparison and high conservation of nucleotide sequences of spa-mxi  
 RT regions between S.sonnei and S.flexneri -- identification of a new  
 RT gene coding plausible membrane protein.";  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: REQUIRED FOR SURFACE PRESENTATION OF INVASION PLASMID  
 CC ANTIGENS. COULD PLAY A ROLE IN PRESERVING THE TRANSLOCATION  
 CC COMPETENCE OF THE IPA ANTIGENS. REQUIRED FOR INVASION AND FOR  
 CC SECRETION OF THE THREE IPA PROTEINS.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: BELONGS TO THE FLIR/MOPE/SPAR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL391753; CAC05830.1; -.  
 DR EMBL; AF348706; AAK18474.1; ALT\_INIT.  
 DR EMBL; D13663; BAA02831.1; -.  
 DR EMBL; AF386526; AAL72306.1; ALT\_INIT.  
 DR EMBL; D50601; BAA09164.1; -.

DR PIR; I49846; I49846.  
 DR InterPro; IPR006304; SpaR\_YscT.  
 DR InterPro; IPR002010; TYPE3IMRPROT.  
 DR Pfam; PF01311; Bac\_export\_1; 1.  
 DR PRINTS; PR00953; TYPE3IMRPROT.  
 DR TIGRFAMs; TIGR01401; fliR\_like\_III; 1.  
 KW Virulence; Transmembrane; Plasmid; Complete proteome.  
 FT TRANSMEM 13 33 POTENTIAL.  
 FT TRANSMEM 37 57 POTENTIAL.  
 FT TRANSMEM 79 99 POTENTIAL.  
 FT TRANSMEM 129 149 POTENTIAL.  
 FT TRANSMEM 183 203 POTENTIAL.  
 FT TRANSMEM 217 237 POTENTIAL.  
 FT VARIANT 168 168 I -> V (IN PLASMIDS PMYSH6000,  
 FT PCP301 AND PLASMID HW383).  
 SQ SEQUENCE 256 AA; 28498 MW; 4B081B38D1FC2A7F CRC64;

Query Match 62.9%; Score 39; DB 1; Length 256;  
 Best Local Similarity 77.8%; Pred. No. 12;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 9  
 ||||| :|  
 Db 27 LFFFLPFLN 35

## RESULT 2

### QAY\_NEUCR

ID QAY\_NEUCR STANDARD; PRT; 537 AA.  
 AC P11636;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Quinate permease (Quinate transporter).  
 GN QA-Y.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=74-OR23-1A / FGSC 987;  
 RX MEDLINE=89293848; PubMed=2525625;  
 RA Geever R.F., Huiet L., Baum J.A., Tyler B.M., Patel V.B.,  
 RA Rutledge B.J., Case M.E., Giles N.H.;  
 RT "DNA sequence, organization and regulation of the qa gene cluster of  
 RT Neurospora crassa.";  
 RL J. Mol. Biol. 207:15-34(1989).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to the sugar transporter family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).

CC

DR EMBL; X14603; CAA32752.1; -.

DR PIR; S04254; G31277.

DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sub\_transporter.

DR InterPro; IPR005829; Sug\_transporter.

DR InterPro; IPR003663; Sugar\_transpt.

DR Pfam; PF00083; sugar\_tr; 1.

DR PRINTS; PR00171; SUGRTRNSPORT.

DR TIGRFAMs; TIGR00879; SP; 1.

DR PROSITE; PS50850; MFS; 1.

DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.

DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.

KW Transmembrane; Transport; Quinate metabolism; Glycoprotein.

FT DOMAIN 1 26 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 27 47 1 (POTENTIAL).

FT DOMAIN 48 74 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 75 95 2 (POTENTIAL).

FT DOMAIN 96 98 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 99 119 3 (POTENTIAL).

FT DOMAIN 120 131 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 132 152 4 (POTENTIAL).

FT DOMAIN 153 160 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 161 181 5 (POTENTIAL).

FT DOMAIN 182 195 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 196 216 6 (POTENTIAL).

FT DOMAIN 217 285 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 286 306 7 (POTENTIAL).

FT DOMAIN 307 327 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 328 349 8 (POTENTIAL).

FT DOMAIN 350 352 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 353 373 9 (POTENTIAL).

FT DOMAIN 374 389 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 390 410 10 (POTENTIAL).

FT DOMAIN 411 435 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 436 456 11 (POTENTIAL).

FT DOMAIN 457 458 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 459 479 12 (POTENTIAL).

FT DOMAIN 480 537 CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 537 AA; 60103 MW; 9AC63400FCC164F3 CRC64;

Query Match 62.9%; Score 39; DB 1; Length 537;

Best Local Similarity 50.0%; Pred. No. 25;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVLP 12

::||||| :|

Db 475 IYFFLPVTKSIP 486

RESULT 3

CNG5\_ARATH

ID CNG5\_ARATH STANDARD; PRT; 717 AA.

AC Q8RWS9; Q9XFS3;

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Probable cyclic nucleotide-gated ion channel 5 (AtCNGC5) (Cyclic  
 DE nucleotide- and calmodulin-regulated ion channel 5).  
 GN CNGC5 OR AT5G57940 OR MTI20.20.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=99272993; PubMed=10341447;  
 RA Koehler C., Merkle T., Neuhaus G.;  
 RT "Characterisation of a novel gene family of putative cyclic  
 RT nucleotide- and calmodulin-regulated ion channels in Arabidopsis  
 RT thaliana.";  
 RL Plant J. 18:97-104(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=98403884; PubMed=9734815;  
 RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.  
 RT Sequence features of the regions of 1,367,185 bp covered by 19  
 RT physically assigned P1 and TAC clones.";  
 RL DNA Res. 5:203-216(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=22954850; PubMed=14593172;  
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,  
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,  
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,  
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Empirical analysis of transcriptional activity in the Arabidopsis  
 RT genome.";  
 RL Science 302:842-846(2003).  
 RN [4]  
 RP GENE FAMILY, AND NOMENCLATURE.  
 RX MEDLINE=21392307; PubMed=11500563;  
 RA Maeser P., Thomine S., Schroeder J.I., Ward J.M., Hirschi K., Sze H.,  
 RA Talke I.N., Amtmann A., Maathuis F.J.M., Sanders D., Harper J.F.,  
 RA Tchieu J., Gribskov M., Persans M.W., Salt D.E., Kim S.A.,  
 RA Guerinot M.L.;  
 RT "Phylogenetic relationships within cation transporter families of  
 RT Arabidopsis.";

```

RL   Plant Physiol. 126:1646-1667(2001).
CC   -!- FUNCTION: Probable cyclic nucleotide-gated ion channel.
CC   -!- SUBUNIT: Homotetramer or heterotetramer (Potential).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane
CC       (Potential).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC       IsoId=Q8RWS9-1; Sequence=Displayed;
CC       Name=2;
CC       IsoId=Q8RWS9-2; Sequence=VSP_008987;
CC       Note=May be due to a competing acceptor splice site. No
CC       experimental confirmation available;
CC   -!- DOMAIN: The binding of calmodulin to the C-terminus might
CC       interfere with cyclic nucleotide binding and thus channel
CC       activation (By similarity).
CC   -!- SIMILARITY: Belongs to the cyclic nucleotide-gated cation channel
CC       (TC 1.A.1.5) family.
CC   -!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
CC   -!- SIMILARITY: Contains 1 IQ domain.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; Y17913; CAB40130.1; -.
DR   EMBL; AB013396; BAB08864.1; -.
DR   EMBL; AY091133; AAM14082.1; -.
DR   EMBL; AY114053; AAM45101.1; -.
DR   PIR; T52573; T52573.
DR   InterPro; IPR000595; cNMP_binding.
DR   InterPro; IPR005821; Ion_trans.
DR   InterPro; IPR000048; IQ_region.
DR   InterPro; IPR001622; K+channel_pore.
DR   Pfam; PF00027; cNMP_binding; 1.
DR   Pfam; PF00520; ion_trans; 1.
DR   Pfam; PF00612; IQ; 1.
DR   SMART; SM00100; cNMP; 1.
DR   PROSITE; PS00888; CNMP_BINDING_1; FALSE_NEG.
DR   PROSITE; PS00889; CNMP_BINDING_2; FALSE_NEG.
DR   PROSITE; PS50042; CNMP_BINDING_3; 1.
DR   PROSITE; PS50096; IQ; 1.
KW   Ion transport; Ionic channel; Calmodulin-binding; cAMP-binding;
KW   cGMP-binding; Transmembrane; Alternative splicing; Multigene family.
FT   DOMAIN          1      102      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM        103     123      H1 (POTENTIAL).
FT   DOMAIN          124     136      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM        137     157      H2 (POTENTIAL).
FT   DOMAIN          158     190      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM        191     211      H3 (POTENTIAL).
FT   DOMAIN          212     224      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM        225     245      H4 (POTENTIAL).
FT   DOMAIN          246     265      CYTOPLASMIC (POTENTIAL).

```

FT	TRANSMEM	266	286	H5 (POTENTIAL).
FT	DOMAIN	287	391	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	392	412	H6 (POTENTIAL).
FT	DOMAIN	413	717	CYTOPLASMIC (POTENTIAL).
FT	NP_BIND	498	628	CNMP.
FT	BINDING	569	569	CAMP OR CGMP (BY SIMILARITY).
FT	DOMAIN	614	629	CALMODULIN-BINDING (BY SIMILARITY).
FT	DOMAIN	634	663	IQ.
FT	VARSPPLIC	10	16	Missing (in isoform 2).
FT				/FTId=VSP_008987.
SQ	SEQUENCE	717 AA;	81968 MW;	444FF45621AE6BDF CRC64;

Query Match 62.9%; Score 39; DB 1; Length 717;  
 Best Local Similarity 75.0%; Pred. No. 32;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFFLPVVN 9  
 ||:||||:  
 Db 117 FFYLPVIN 124

#### RESULT 4

##### Y3H1\_ANASP

ID Y3H1\_ANASP STANDARD; PRT; 364 AA.  
 AC Q8YQ64;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical zinc metalloprotease All13971 (EC 3.4.24.-).  
 GN ALL13971.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120."  
 RL DNA Res. 8:205-213(2001).  
 CC -!- COFACTOR: Zinc (Probable).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (By similarity).  
 CC -!- SIMILARITY: Belongs to peptidase family M50B.  
 CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AP003594; BAB75670.1; -.  
 DR PIR; AD2302; AD2302.  
 DR MEROPS; M50.004; -.  
 DR InterPro; IPR001478; PDZ.  
 DR InterPro; IPR004387; Pept\_M50\_Zn.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR008915; Peptidase\_M50.  
 DR Pfam; PF00595; PDZ; 1.  
 DR Pfam; PF02163; Peptidase\_M50; 1.  
 DR SMART; SM00228; PDZ; 1.  
 DR TIGRFAMs; TIGR00054; TIGR00054; 1.  
 DR PROSITE; PS50106; PDZ; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Hypothetical protein; Hydrolase; Metalloprotease; Zinc; Transmembrane;  
 KW Inner membrane; Complete proteome.  
 FT METAL 17 17 ZINC (CATALYTIC) (POTENTIAL).  
 FT ACT\_SITE 18 18 POTENTIAL.  
 FT METAL 21 21 ZINC (CATALYTIC) (POTENTIAL).  
 FT TRANSMEM 92 114 POTENTIAL.  
 FT TRANSMEM 281 303 POTENTIAL.  
 FT TRANSMEM 329 346 POTENTIAL.  
 FT DOMAIN 103 188 PDZ.  
 SQ SEQUENCE 364 AA; 38613 MW; 54F6AAE818AEFBEA CRC64;

Query Match 62.1%; Score 38.5; DB 1; Length 364;  
 Best Local Similarity 47.4%; Pred. No. 21;  
 Matches 9; Conservative 2; Mismatches 1; Indels 7; Gaps 1;

Qy 1 LFFF-----LPVVNVLP 12  
 |||| | |::||  
 Db 281 LFFFAALISINLAVINILP 299

# RESULT 5

## YCJJ\_ECOLI

ID YCJJ\_ECOLI STANDARD; PRT; 461 AA.  
 AC P76037; P77557;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical transport protein ycjJ.  
 GN YCJJ OR B1296.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12."  
 RL Science 277:1453-1474(1997).



RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97251357; PubMed=9097039;  
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,  
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,  
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
 RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,  
 RA Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,  
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;  
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 28.0-40.1 min region on the linkage map.";  
 RL DNA Res. 3:363-377(1996).  
 CC -!- FUNCTION: Probable amino-acid or metabolite transport protein.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -!- SIMILARITY: Belongs to the amino acid permease family.  
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to a  
 CC frameshift in position 10.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000227; AAC74378.1; ALT\_INIT.  
 DR EMBL; D90767; BAA14856.1; ALT\_FRAME.  
 DR EMBL; D90768; BAA14865.1; ALT\_FRAME.  
 DR EcoGene; EG13907; ycjJ.  
 DR InterPro; IPR002293; AA/rel\_permease1.  
 DR InterPro; IPR004840; AAc\_permease.  
 DR InterPro; IPR004841; Permease\_region.  
 DR Pfam; PF00324; aa\_permeases; 1.  
 DR PROSITE; PS00218; AMINO\_ACID\_PERMEASE\_1; FALSE\_NEG.  
 KW Hypothetical protein; Transport; Amino-acid transport; Transmembrane;  
 KW Inner membrane; Complete proteome.  
 FT TRANSMEM 27 47 POTENTIAL.  
 FT TRANSMEM 54 74 POTENTIAL.  
 FT TRANSMEM 100 120 POTENTIAL.  
 FT TRANSMEM 125 145 POTENTIAL.  
 FT TRANSMEM 162 182 POTENTIAL.  
 FT TRANSMEM 204 224 POTENTIAL.  
 FT TRANSMEM 242 262 POTENTIAL.  
 FT TRANSMEM 275 295 POTENTIAL.  
 FT TRANSMEM 344 364 POTENTIAL.  
 FT TRANSMEM 365 385 POTENTIAL.  
 FT TRANSMEM 399 419 POTENTIAL.  
 FT TRANSMEM 422 442 POTENTIAL.  
 SQ SEQUENCE 461 AA; 50853 MW; B027F1FA01C1B5BB CRC64;

Query Match 61.3%; Score 38; DB 1; Length 461;  
 Best Local Similarity 60.0%; Pred. No. 32;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FFFLPVVNVL 11  
: |||::|||  
Db 110 YLFLPMINVL 119

RESULT 6

CG15\_ARATH

ID CG15\_ARATH STANDARD; PRT; 678 AA.

AC Q9SL29;

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Putative cyclic nucleotide-gated ion channel 15 (Cyclic nucleotide-  
DE and calmodulin-regulated ion channel 15).

GN CNGC15 OR AT2G28260 OR T3B23.7.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RX MEDLINE=20083487; PubMed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,

RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,

RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,

RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,

RA Goodman H.M., Somerville C.R., Copenhagen G.P., Preuss D.,

RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,

RA Venter J.C.;

RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis

RT thaliana.";

RL Nature 402:761-768(1999).

RN [2]

RP GENE FAMILY, AND NOMENCLATURE.

RX MEDLINE=21392307; PubMed=11500563;

RA Maeser P., Thomine S., Schroeder J.I., Ward J.M., Hirschi K., Sze H.,

RA Talke I.N., Amtmann A., Maathuis F.J.M., Sanders D., Harper J.F.,

RA Tchieu J., Gribskov M., Persans M.W., Salt D.E., Kim S.A.,

RA Guerinot M.L.;

RT "Phylogenetic relationships within cation transporter families of

RT Arabidopsis.";

RL Plant Physiol. 126:1646-1667(2001).

CC -!- FUNCTION: Putative cyclic nucleotide-gated ion channel.

CC -!- SUBUNIT: Homotetramer or heterotetramer (Potential).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane  
CC (Potential).

CC -!- DOMAIN: The binding of calmodulin to the C-terminus might  
CC interfere with cyclic nucleotide binding and thus channel  
CC activation (By similarity).

CC -!- SIMILARITY: Belongs to the cyclic nucleotide-gated cation channel  
CC (TC 1.A.1.5) family.

CC -!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.

CC -!- SIMILARITY: Contains 1 IQ domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AC006202; AAD29827.1; -.  
DR PIR; G84682; G84682.  
DR InterPro; IPR000595; cNMP\_binding.  
DR InterPro; IPR005821; Ion\_trans.  
DR InterPro; IPR000048; IQ\_region.  
DR InterPro; IPR001622; K+channel\_pore.  
DR Pfam; PF00027; cNMP\_binding; 1.  
DR Pfam; PF00520; ion\_trans; 1.  
DR Pfam; PF00612; IQ; 1.  
DR SMART; SM00100; cNMP; 1.  
DR PROSITE; PS00888; CNMP\_BINDING\_1; FALSE\_NEG.  
DR PROSITE; PS00889; CNMP\_BINDING\_2; FALSE\_NEG.  
DR PROSITE; PS50042; CNMP\_BINDING\_3; 1.  
DR PROSITE; PS50096; IQ; 1.  
KW Hypothetical protein; Ion transport; Ionic channel;  
KW Calmodulin-binding; cAMP-binding; cGMP-binding; Transmembrane;  
KW Multigene family.  
FT DOMAIN 1 81 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 82 102 H1 (POTENTIAL).  
FT DOMAIN 103 115 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 116 136 H2 (POTENTIAL).  
FT DOMAIN 137 170 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 171 191 H3 (POTENTIAL).  
FT DOMAIN 192 203 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 204 224 H4 (POTENTIAL).  
FT DOMAIN 225 245 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 246 266 H5 (POTENTIAL).  
FT DOMAIN 267 364 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 365 385 H6 (POTENTIAL).  
FT DOMAIN 386 678 CYTOPLASMIC (POTENTIAL).  
FT NP\_BIND 471 595 CNMP.  
FT BINDING 542 542 CAMP OR CGMP (BY SIMILARITY).  
FT DOMAIN 587 602 CALMODULIN-BINDING (BY SIMILARITY).  
FT DOMAIN 607 638 IQ.  
SQ SEQUENCE 678 AA; 78722 MW; E020D14E44050E64 CRC64;

Query Match 61.3%; Score 38; DB 1; Length 678;  
Best Local Similarity 87.5%; Pred. No. 46;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVV 8  
| | | | | | | :  
Db 96 LFFFLPVM 103

RESULT 7  
CNG7\_ARATH  
ID CNG7\_ARATH STANDARD; PRT; 738 AA.  
AC Q9S9N5;

DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Putative cyclic nucleotide-gated ion channel 7 (Cyclic nucleotide- and  
 DE calmodulin-regulated ion channel 7).  
 GN CNGC7 OR AT1G15990 OR T24D18.9.  
 OS *Arabidopsis thaliana* (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; *Arabidopsis*.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,  
 RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant *Arabidopsis*  
 RT *thaliana*.";  
 RL Nature 408:816-820(2000).  
 RN [2]  
 RP GENE FAMILY, AND NOMENCLATURE.  
 RX MEDLINE=21392307; PubMed=11500563;  
 RA Maeser P., Thomine S., Schroeder J.I., Ward J.M., Hirschi K., Sze H.,  
 RA Talke I.N., Amtmann A., Maathuis F.J.M., Sanders D., Harper J.F.,  
 RA Tchieu J., Gribskov M., Persans M.W., Salt D.E., Kim S.A.,  
 RA Gueriot M.L.;  
 RT "Phylogenetic relationships within cation transporter families of  
 RT *Arabidopsis*.";  
 RL Plant Physiol. 126:1646-1667(2001).  
 CC -!- FUNCTION: Putative cyclic nucleotide-gated ion channel.  
 CC -!- SUBUNIT: Homotetramer or heterotetramer (Potential).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane  
 CC (Potential).  
 CC -!- DOMAIN: The binding of calmodulin to the C-terminus might  
 CC interfere with cyclic nucleotide binding and thus channel  
 CC activation (By similarity).  
 CC -!- SIMILARITY: Belongs to the cyclic nucleotide-gated cation channel  
 CC (TC 1.A.1.5) family.  
 CC -!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.  
 CC -!- SIMILARITY: Contains 1 IQ domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AC010924; AAF18496.1; -.  
 DR PIR; E86294; E86294.  
 DR InterPro; IPR000595; cNMP\_binding.  
 DR InterPro; IPR005821; Ion\_trans.  
 DR InterPro; IPR000048; IQ\_region.  
 DR InterPro; IPR001622; K+channel\_pore.  
 DR Pfam; PF00027; cNMP\_binding; 1.  
 DR Pfam; PF00520; ion\_trans; 1.  
 DR Pfam; PF00612; IQ; 1.  
 DR SMART; SM00100; cNMP; 1.  
 DR PROSITE; PS00888; CNMP\_BINDING\_1; FALSE\_NEG.  
 DR PROSITE; PS00889; CNMP\_BINDING\_2; FALSE\_NEG.  
 DR PROSITE; PS50042; CNMP\_BINDING\_3; 1.  
 DR PROSITE; PS50096; IQ; FALSE\_NEG.  
 KW Hypothetical protein; Ion transport; Ionic channel;  
 KW Calmodulin-binding; cAMP-binding; cGMP-binding; Transmembrane;  
 KW Multigene family.  
 FT DOMAIN 1 104 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 105 125 H1 (POTENTIAL).  
 FT DOMAIN 126 139 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 140 160 H2 (POTENTIAL).  
 FT DOMAIN 161 193 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 194 214 H3 (POTENTIAL).  
 FT DOMAIN 215 227 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 228 248 H4 (POTENTIAL).  
 FT DOMAIN 249 268 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 269 289 H5 (POTENTIAL).  
 FT DOMAIN 290 395 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 396 416 H6 (POTENTIAL).  
 FT DOMAIN 417 738 CYTOPLASMIC (POTENTIAL).  
 FT NP\_BIND 502 632 CNMP.  
 FT BINDING 573 573 CAMP OR CGMP (BY SIMILARITY).  
 FT DOMAIN 618 633 CALMODULIN-BINDING (BY SIMILARITY).  
 FT DOMAIN 638 667 IQ.  
 SQ SEQUENCE 738 AA; 84634 MW; 369AB538E25959BF CRC64;

Query Match 61.3%; Score 38; DB 1; Length 738;  
 Best Local Similarity 66.7%; Pred. No. 50;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 9  
 |||:|:|:  
 Db 118 LFFYLPIVD 126

RESULT 8  
 CNG8\_ARATH  
 ID CNG8\_ARATH STANDARD; PRT; 753 AA.  
 AC Q9FXH6;  
 DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Putative cyclic nucleotide-gated ion channel 8 (Cyclic nucleotide- and  
 DE calmodulin-regulated ion channel 8).  
 GN CNGC8 OR AT1G19780 OR F6F9.17 OR F14P1.12.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,  
 RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 408:816-820(2000).  
 RN [2]  
 RP CONCEPTUAL TRANSLATION.  
 RA Tognolli M.;  
 RL Unpublished observations (OCT-2003).  
 RN [3]  
 RP GENE FAMILY, AND NOMENCLATURE.  
 RX MEDLINE=21392307; PubMed=11500563;  
 RA Maeser P., Thomine S., Schroeder J.I., Ward J.M., Hirschi K., Sze H.,  
 RA Talke I.N., Amtmann A., Maathuis F.J.M., Sanders D., Harper J.F.,  
 RA Tchieu J., Gribskov M., Persans M.W., Salt D.E., Kim S.A.,  
 RA Gueriot M.L.;  
 RT "Phylogenetic relationships within cation transporter families of  
 RT Arabidopsis.";  
 RL Plant Physiol. 126:1646-1667(2001).  
 CC -!- FUNCTION: Putative cyclic nucleotide-gated ion channel.  
 CC -!- SUBUNIT: Homotetramer or heterotetramer (Potential).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Plasma membrane  
 CC (Potential).  
 CC -!- DOMAIN: The binding of calmodulin to the C-terminus might  
 CC interfere with cyclic nucleotide binding and thus channel  
 CC activation (By similarity).  
 CC -!- SIMILARITY: Belongs to the cyclic nucleotide-gated cation channel  
 CC (TC 1.A.1.5) family.  
 CC -!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.

CC -!- SIMILARITY: Contains 1 IQ domain.  
CC -!- CAUTION: Ref.1 sequence differs from that shown due to erroneous  
CC gene model prediction.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AC007797; AAG12561.1; ALT\_SEQ.  
DR EMBL; AC024609; -; NOT\_ANNOTATED\_CDS.  
DR PIR; H86330; H86330.  
DR InterPro; IPR000595; cNMP\_binding.  
DR InterPro; IPR005821; Ion\_trans.  
DR InterPro; IPR000048; IQ\_region.  
DR Pfam; PF00027; cNMP\_binding; 1.  
DR Pfam; PF00520; ion\_trans; 1.  
DR Pfam; PF00612; IQ; 1.  
DR SMART; SM00100; cNMP; 1.  
DR PROSITE; PS00888; CNMP\_BINDING\_1; FALSE\_NEG.  
DR PROSITE; PS00889; CNMP\_BINDING\_2; FALSE\_NEG.  
DR PROSITE; PS50042; CNMP\_BINDING\_3; 1.  
DR PROSITE; PS50096; IQ; 1.  
KW Hypothetical protein; Ion transport; Ionic channel;  
KW Calmodulin-binding; cAMP-binding; cGMP-binding; Transmembrane;  
KW Multigene family.  
FT DOMAIN 1 111 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 112 132 H1 (POTENTIAL).  
FT DOMAIN 133 145 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 146 166 H2 (POTENTIAL).  
FT DOMAIN 167 199 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 200 220 H3 (POTENTIAL).  
FT DOMAIN 221 233 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 234 254 H4 (POTENTIAL).  
FT DOMAIN 255 274 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 275 295 H5 (POTENTIAL).  
FT DOMAIN 296 402 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 403 423 H6 (POTENTIAL).  
FT DOMAIN 424 753 CYTOPLASMIC (POTENTIAL).  
FT NP\_BIND 508 638 CNMP.  
FT BINDING 579 579 CAMP OR CGMP (BY SIMILARITY).  
FT DOMAIN 624 639 CALMODULIN-BINDING (BY SIMILARITY).  
FT DOMAIN 644 673 IQ.  
SQ SEQUENCE 753 AA; 85982 MW; 5AB2A74E25BC7CE4 CRC64;

Query Match 61.3%; Score 38; DB 1; Length 753;  
Best Local Similarity 66.7%; Pred. No. 51;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 9  
|||:|:|:  
Db 125 LFFYLPIVD 133

RESULT 9

SPAR\_SALTY

ID SPAR\_SALTY STANDARD; PRT; 263 AA.  
AC P40701;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Surface presentation of antigens protein spaR.  
GN SPAR OR STM2888.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94008985; PubMed=8404849;  
RA Groisman E.A., Ochman H.;  
RT "Cognate gene clusters govern invasion of host epithelial cells by  
RT Salmonella typhimurium and Shigella flexneri.";  
RL EMBO J. 12:3779-3787(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
RX MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
RT LT2.";  
RL Nature 413:852-856(2001).  
CC -!- FUNCTION: INVOLVED IN A SECRETORY PATHWAY RESPONSIBLE FOR THE  
CC SURFACE PRESENTATION OF DETERMINANTS NEEDED FOR THE ENTRY OF  
CC SALMONELLA SPECIES INTO MAMMALIAN CELLS.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -!- SIMILARITY: BELONGS TO THE FLIR/MOPE/SPAR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X73525; CAA51926.1; -.  
DR EMBL; AE008832; AAL21768.1; -.  
DR PIR; S37309; S37309.  
DR StyGene; SG10470; spaR.  
DR InterPro; IPR006304; SpaR\_YscT.  
DR InterPro; IPR002010; TYPE3IMRPROT.  
DR Pfam; PF01311; Bac\_export\_1; 1.  
DR PRINTS; PR00953; TYPE3IMRPROT.  
DR TIGRFAMs; TIGR01401; flir\_like\_III; 1.  
KW Virulence; Transmembrane; Complete proteome.  
FT TRANSMEM 12 32 POTENTIAL.



FT TRANSMEM 46 66 POTENTIAL.  
 FT TRANSMEM 82 102 POTENTIAL.  
 FT TRANSMEM 127 147 POTENTIAL.  
 FT TRANSMEM 182 202 POTENTIAL.  
 FT TRANSMEM 211 231 POTENTIAL.  
 SQ SEQUENCE 263 AA; 28486 MW; B267D1EB643DA3C5 CRC64;

Query Match 59.7%; Score 37; DB 1; Length 263;  
 Best Local Similarity 66.7%; Pred. No. 28;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVVN 9  
 :||||| :|  
 Db 26 IFFFLPFLN 34

# RESULT 10

## PPCK\_PSESM

ID PPCK\_PSESM STANDARD; PRT; 514 AA.  
 AC Q88AZ4;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49) (PEP  
 DE carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).  
 GN PCKA OR PSPTO0239.  
 OS Pseudomonas syringae (pv. tomato).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=323;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DC3000;  
 RX MEDLINE=22834015; PubMed=12928499;  
 RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,  
 RA Gwinn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,  
 RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,  
 RA Nelson W.C., Davidsen T., Zafar N., Zhou L., Liu J., Yuan Q.,  
 RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,  
 RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,  
 RA Alfano J.R., Cartinhour S., Chatterjee A.K., Delaney T.P.,  
 RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,  
 RA White O., Fraser C.M., Collmer A.;  
 RT "The complete genome sequence of the Arabidopsis and tomato pathogen  
 RT Pseudomonas syringae pv. tomato DC3000.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).  
 CC -!- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate  
 CC + CO(2).  
 CC -!- PATHWAY: Rate-limiting gluconeogenic enzyme.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [ATP]  
 CC family.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AE016856; AAO53785.1; -.  
DR TIGR; PSPT00239; -.  
DR HAMAP; MF\_00453; -; 1.  
DR InterPro; IPR001272; PEPCK\_ATP.  
DR Pfam; PF01293; PEPCK\_ATP; 1.  
DR PROSITE; PS00532; PEPCK\_ATP; FALSE\_NEG.  
KW Gluconeogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.  
FT NP\_BIND 220 227 ATP (BY SIMILARITY).  
SQ SEQUENCE 514 AA; 55809 MW; 09DB88AD11CC01D3 CRC64;

Query Match 59.7%; Score 37; DB 1; Length 514;  
Best Local Similarity 70.0%; Pred. No. 53;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 FFLPVVNVLP 12  
| || |:|  
Db 193 FLLPAVDVLP 202

RESULT 11

LNT\_TREPA

ID LNT\_TREPA STANDARD; PRT; 546 AA.  
AC O83279;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Apolipoprotein N-acyltransferase (EC 2.3.1.-) (ALP N-acyltransferase).  
GN LNT OR TP0252.  
OS Treponema pallidum.  
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
OX NCBI\_TaxID=160;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Nichols;  
RX MEDLINE=98332770; PubMed=9665876;  
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,  
RA McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
RA Venter J.C.;  
RT "Complete genome sequence of Treponema pallidum, the syphilis  
RT spirochete."  
RL Science 281:375-388(1998).  
CC -!- FUNCTION: Transfers the fatty acyl group on membrane lipoproteins  
CC (By similarity).  
CC -!- PATHWAY: Lipoproteins biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
CC -!- SIMILARITY: Belongs to the apolipoprotein N-acyltransferase  
CC family.  
CC -!- SIMILARITY: Contains 1 CN hydrolase domain.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE001206; AAC65237.1; -.  
 DR PIR; G71348; G71348.  
 DR TIGR; TP0252; -.  
 DR InterPro; IPR003010; Ntlse/CNhydrtse.  
 DR Pfam; PF00795; CN\_hydrolase; 1.  
 DR PROSITE; PS50263; CN\_HYDROLASE; 1.  
 KW Transferase; Acyltransferase; Transmembrane; Complete proteome.  
 FT TRANSMEM 14 34 POTENTIAL.  
 FT TRANSMEM 62 82 POTENTIAL.  
 FT TRANSMEM 85 105 POTENTIAL.  
 FT TRANSMEM 122 142 POTENTIAL.  
 FT TRANSMEM 151 171 POTENTIAL.  
 FT TRANSMEM 194 214 POTENTIAL.  
 FT TRANSMEM 490 510 POTENTIAL.  
 FT TRANSMEM 514 534 POTENTIAL.  
 FT DOMAIN 233 546 CN HYDROLASE.  
 SQ SEQUENCE 546 AA; 61513 MW; 06E8041A3FB8821E CRC64;

Query Match 59.7%; Score 37; DB 1; Length 546;  
 Best Local Similarity 77.8%; Pred. No. 56;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 FLPVVNVLP 12  
 |:| |||||  
 Db 345 FIPGVNVLP 353

# RESULT 12

## NKX1\_CHICK

ID NKX1\_CHICK STANDARD; PRT; 663 AA.  
 AC Q9IAL8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Sodium/potassium/calcium exchanger 1 precursor (Na(+)/K(+)/Ca(2+)-  
 DE exchange protein 1) (Retinal rod Na-Ca+K exchanger).  
 GN SLC24A1 OR NCKX1.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Retina;  
 RX MEDLINE=20130359; PubMed=10662833;  
 RA Prinsen C.F.M., Szerencsei R.T., Schnetkamp P.P.M.;  
 RT "Molecular cloning and functional expression of the potassium-  
 RT dependent sodium-calcium exchanger from human and chicken retinal cone

RT photoreceptors.";  
 RL J. Neurosci. 20:1424-1434(2000).  
 CC -!- FUNCTION: Critical component of the visual transduction cascade,  
 CC controlling the calcium concentration of outer segments during  
 CC light and darkness. Light causes a rapid lowering of cytosolic  
 CC free calcium in the outer segment of both retinal rod and cone  
 CC photoreceptors and the light-induced lowering of calcium is caused  
 CC by extrusion via this protein which plays a key role in the  
 CC process of light adaptation. Transports one Ca(2+) and one K(+) in  
 CC exchange for four Na(+).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Retinal rods. Localizes to the inner segment  
 CC of rod photoreceptors.  
 CC -!- SIMILARITY: BELONGS TO THE SLC24A FAMILY OF TRANSPORTERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF177984; AAF25808.1; -.  
 DR InterPro; IPR004817; K\_NaCaexchang.  
 DR InterPro; IPR004481; K\_NaCaexchnng.  
 DR InterPro; IPR004837; NaCa\_Exmemb.  
 DR Pfam; PF01699; Na\_Ca\_Ex; 2.  
 DR TIGRFAMs; TIGR00927; 2A1904; 1.  
 DR TIGRFAMs; TIGR00367; TIGR00367; 1.  
 KW Vision; Transport; Antiport; Symport; Calcium transport;  
 KW Potassium transport; Sodium transport; Transmembrane; Glycoprotein;  
 KW Phosphorylation; Signal; Repeat.  
 FT SIGNAL 1 31 POTENTIAL.  
 FT CHAIN 32 663 SODIUM/POTASSIUM/CALCIUM EXCHANGER 1.  
 FT DOMAIN 32 128 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 129 149 POTENTIAL.  
 FT DOMAIN 150 173 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 174 194 POTENTIAL.  
 FT DOMAIN 195 200 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 201 221 POTENTIAL.  
 FT DOMAIN 222 228 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 229 253 POTENTIAL.  
 FT DOMAIN 254 259 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 260 276 POTENTIAL.  
 FT DOMAIN 277 471 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 472 492 POTENTIAL.  
 FT DOMAIN 493 499 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 500 520 POTENTIAL.  
 FT DOMAIN 521 535 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 536 556 POTENTIAL.  
 FT DOMAIN 557 574 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 575 595 POTENTIAL.  
 FT DOMAIN 596 604 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 605 625 POTENTIAL.  
 FT DOMAIN 626 632 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 633 653 POTENTIAL.

FT	DOMAIN	654	663	CYTOPLASMIC (POTENTIAL).
FT	REPEAT	170	210	ALPHA-1.
FT	REPEAT	543	574	ALPHA-2.
FT	MOD_RES	337	337	PHOSPHORYLATION (POTENTIAL).
FT	CARBOHYD	59	59	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	66	66	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	100	100	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	663 AA; 73771 MW; DD624E3080C43082 CRC64;		

Query Match 59.7%; Score 37; DB 1; Length 663;  
 Best Local Similarity 63.6%; Pred. No. 67;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVL 11  
 :|||| ||::|  
 Db 14 IFFFLAVVSLL 24

# RESULT 13

YD68\_METJA

ID YD68 METJA STANDARD; PRT; 249 AA.  
 AC Q58763;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Putative ABC transporter permease protein MJ1368.  
 GN MJ1368.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii.";  
 RL Science 273:1058-1073(1996).  
 CC -!- FUNCTION: Might be part of an ABC transporter complex. Might be  
 CC responsible for the translocation of the substrate across the  
 CC membrane.  
 CC -!- SUBUNIT: Might form a complex with the ATP-binding protein MJ1367.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U67576; AAB99376.1; -.  
DR PIR; G64470; G64470.  
DR TIGR; MJ1368; -.  
DR InterPro; IPR000515; BPD\_transp.  
DR Pfam; PF00528; BPD\_transp; 1.  
DR PROSITE; PS50928; ABC\_TM1; 1.  
KW Hypothetical protein; Transport; Transmembrane; Complete proteome.  
FT TRANSMEM 6 26 POTENTIAL.  
FT TRANSMEM 50 70 POTENTIAL.  
FT TRANSMEM 94 114 POTENTIAL.  
FT TRANSMEM 117 137 POTENTIAL.  
FT TRANSMEM 179 199 POTENTIAL.  
FT TRANSMEM 224 244 POTENTIAL.  
SQ SEQUENCE 249 AA; 27941 MW; 3359BC2A3EB48675 CRC64;

Query Match 58.1%; Score 36; DB 1; Length 249;  
Best Local Similarity 54.5%; Pred. No. 39;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LFFFLPVVNVL 11  
|| |||:::|  
Db 17 LFIFLPIIYML 27

#### RESULT 14

##### MRAY\_THEMA

ID MRAY\_THEMA STANDARD; PRT; 302 AA.  
AC Q9WY77;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-  
DE MurNAc-pentapeptide phosphotransferase).  
GN MRAY OR TM0235.  
OS Thermotoga maritima.  
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.  
OX NCBI\_TaxID=2336;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;  
RX MEDLINE=99287316; PubMed=10360571;  
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,  
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
RT genome sequence of Thermotoga maritima."  
RL Nature 399:323-329(1999).  
CC -!- FUNCTION: First step of the lipid cycle reactions in the  
CC biosynthesis of the cell wall peptidoglycan (By similarity).  
CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-  
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-

```

CC      Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC      -!- PATHWAY: Peptidoglycan biosynthesis.
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC      -!- SIMILARITY: Belongs to the glycosyltransferase family 4. MraY
CC      subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AE001707; AAD35326.1; -.
DR      PIR; E72402; E72402.
DR      TIGR; TM0235; -.
DR      HAMAP; MF_00038; -, 1.
DR      InterPro; IPR000715; Glyco_trans_4.
DR      InterPro; IPR003524; PNAcPpept_trans.
DR      Pfam; PF00953; Glycos_transf_4; 1.
DR      TIGRFAMs; TIGR00445; mraY; 1.
DR      PROSITE; PS01347; MRAY_1; 1.
DR      PROSITE; PS01348; MRAY_2; 1.
KW      Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
KW      Complete proteome.
FT      TRANSMEM      1      21      POTENTIAL.
FT      TRANSMEM      42      62      POTENTIAL.
FT      TRANSMEM      67      87      POTENTIAL.
FT      TRANSMEM      95     115      POTENTIAL.
FT      TRANSMEM     123     143      POTENTIAL.
FT      TRANSMEM     154     174      POTENTIAL.
FT      TRANSMEM     178     198      POTENTIAL.
FT      TRANSMEM     204     224      POTENTIAL.
FT      TRANSMEM     229     249      POTENTIAL.
FT      TRANSMEM     281     301      POTENTIAL.
SQ      SEQUENCE      302 AA;  33814 MW;  BB8FF74FEA9205CB CRC64;

      Query Match          58.1%;  Score 36;  DB 1;  Length 302;
      Best Local Similarity 54.5%;  Pred. No. 47;
      Matches      6;  Conservative      2;  Mismatches      3;  Indels      0;  Gaps      0;

Qy      1 LFFFLPVVNVL 11
      ||| :||: |
Db      233 LFFMIPVIETL 243

```

# RESULT 15

YTR1\_BUCSC

```

ID      YTR1_BUCSC      STANDARD;      PRT;      304 AA.
AC      Q44601;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Hypothetical transport protein in trpA 3'region.
OS      Buchnera aphidicola (subsp. Schlechtendalia chinensis).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

OC Enterobacteriaceae; Buchnera.  
 OX NCBI\_TaxID=118110;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95261545; PubMed=7742976;  
 RA Lai C.-Y., Baumann P., Moran N.A.;  
 RT "Genetics of the tryptophan biosynthetic pathway of the prokaryotic  
 endosymbiont (Buchnera) of the aphid *Schlechtendalia chinensis*.";  
 RL Insect Mol. Biol. 4:47-59(1995).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -!- SIMILARITY: BELONGS TO THE EAMA TRANSPORTER FAMILY. STRONG, TO  
 CC S.TYPHIMURIUM PAGO.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U09185; AAA92793.1; -.  
 DR InterPro; IPR000620; DUF6.  
 DR Pfam; PF00892; DUF6; 2.  
 KW Hypothetical protein; Transport; Transmembrane.  
 FT TRANSMEM 5 25 POTENTIAL.  
 FT TRANSMEM 42 62 POTENTIAL.  
 FT TRANSMEM 68 88 POTENTIAL.  
 FT TRANSMEM 96 116 POTENTIAL.  
 FT TRANSMEM 120 140 POTENTIAL.  
 FT TRANSMEM 150 170 POTENTIAL.  
 FT TRANSMEM 178 198 POTENTIAL.  
 FT TRANSMEM 215 235 POTENTIAL.  
 FT TRANSMEM 245 265 POTENTIAL.  
 FT TRANSMEM 268 288 POTENTIAL.  
 SQ SEQUENCE 304 AA; 35051 MW; A6676E9610E55EAE CRC64;  
  
 Query Match 58.1%; Score 36; DB 1; Length 304;  
 Best Local Similarity 36.4%; Pred. No. 47;  
 Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
  
 Qy 1 LFFFLPVVNVL 11  
 :|| |::|::  
 Db 248 IFFIFPIINLM 258

Search completed: August 24, 2004, 15:43:28  
 Job time : 9.44776 secs